POLSKA AKADEMIA NAUK
INSTYTUT FIZYKI

ACTA PHYSICA POLONICA

DWUMIESIĘCZNIK

Vol. XIX — Fasc. 4

WARSZAWA 1960

First reprint by PWN — Polish Scientific Publishers 1969

ON THE FLUORESCENCE OF Sb- AND Mn-ACTIVATED CALCIUM HALOPHOSPHATES

By E. Ostaszewicz

Institute of General Physics C, Faculty of Chemistry, Polytechnical High School, Warsaw

(Received May 8, 1959)

Sb- and Mn-activated calcium halophosphates were produced and their fluorescence was investigated by the spectral method in its dependence on the activator concentration and the weight ratio of the halogens. The trichromatic colour coordinates were computed from the experimental intensity distribution. The maximum quantum yield of the fluorescence as a function of the concentration of activators and halogens and a displacement of the fluorescence maximum of Mn was found to occur in the spectral range of 5760—5935 Å when fluorine had been entirely replaced by chlorine. The fluorescence mechanism is discussed in detail; the one consiting in direct resonance energy transfer from the sensitizer to the activator was found to agree best with the experimental results. According to the author's results, the range of the Sb to Mn energy transfer amounts to 50 cation loci. Perturbation theory as applied to resonance transfer of energy in solids yielded an explanation of the possibility of Mn transitions into forbidden states through the agency of the sensitizer, the return wherefrom gives rise to the orange fluorescence characteristic of Mn.

Introduction

Calcium halophosphates were discovered by Mc Keag and Ranby in 1942 in Great Britain (1942). The basic substance of formula $3Ca_3(PO_4)_2\,Ca(FCl)_2$ exhibits apatite structure; the activators are antimony and manganium. The compounds are intensely fluorescent when excited with the 2537 Å mercury resonance line, and hence are widely applied in lighting techniques in low-pressure mercury lamps. They are of a high degree of interest to scientific research, too, as belonging to the group of doubly activated luminophors, the discovery of which occurred at a rather late period.

For this reason, the compounds have been the object of investigation by various authors. At first, principally the technology of production of calcium holophosphates was dealt with; later investigation centered on the mechanism of luminescence of the luminophors.

Jenkins and co-workers (1949) elaborated a technological method of obtaining calcium halophosphates, and investigated the fluorescence of the latter in its depen-

dence on the concentration of the activators and on the atomic weight of the halogen used.

Further, extensive work on calcium halophosphates was carried out by Butler and Jerome (1950).

Bandel and co-workers (1951) investigated the properties of the fluorescence of calcium halophosphates as dependent on the chemical composition of the luminophor, the weight ratio of the halogens, and the amount of activators introduced.

Doherty and Harrison (1954) worked out an extensive method of preparing the salts composing calcium halophosphates, and investigated the fluorescence of a number of luminophors obtained therefrom.

It was the aim of the present paper to obtain a variety of calcium halophosphates, to investigate their fluorescence spectrum by the spectral method as dependent on the concentration of the activators and on the quantitative ratio of the halogens, to compare the results with those obtained previously, and to determine the mechanism of light emission.

1. Technology of Production of Calcium Halophosphates

The present author obtained calcium halophosphates by the method of sintering mixtures of the following salts:

1. $CaHPO_4$
2. $CaCO_3$
3. $(NH_4)_2\,HPO_4$
4. CaF_2
5. $CaCl_2$
6. Sb_2O_3
7. $Mn_3\,(PO_4)_2$

These were obtained from the Production Centre of Chemical Reagents (Wytwórnia Odczynników Chemicznych) at Gliwice, Poland, with the designation as being „of special purity for halophosphates". Previous to sintering, the salts were throughly dried at about 300°C and ground to a grain size of 5 — 10 μ. On mixing the components in the appropriate wieght ratio, the mixture was sintered during 35 minutes in an open quartz tube at 1050° to 1110°C. Mixtures containing less Mn required a lower sintering temperature, whereas those with a higher Mn content required a higher temperature. Previous to introducing the tube into the oven, the latter was heated to the temperature required. Sintering proceeded in an atmosphere of the gases leaving the tube, as: CO_2, P_2O_5, NH_3, N_2 and $SbCl_3$. The technology of the sintering process is known to consist on the following: the high temperature leads to dissociation of the different molecules of the mixture into ions, which then regroup to form crystals of calcium halophosphate with ions of the activators incorporated. The remaining ions combine to gas molecules and leave the site of the reaction, vola-

tilizing process, from 22 to 24 % of the mixture volatilized. It is required that the atmosphere of sintering should not have an oxidating effect, as the activators can incorporate into crystal lattice of the calcium halophosphate in the form of ions of the lowest valences. This requirement is set by the ionic radii. The product, on cooling in an anhydrous atmosphere of very low oxygen content, was powdered in a mortar, thus yielding the ready luminophor. In this manner the author prepared two kinds of luminophors: one containing luminophors of the constant fluorine to chlorine weight ratio of 49 : 51 and a variable weight ratio of the activators, and another consisting of luminophors of a constant weight ratio of the activators and a variable ratio of the halogens. The foregoing amounts of activators and halogens in weight per cent relate to the mixtures previous to sintering. Since Sb and Cl_2 volatilize in the process of sintering in the form of $SbCl_3$, the respective figures serve primarily to provide orientation as to their amounts present in the luminophors.[1])

2. Account of Experimental Method and Apparature Used

The fluorescence was investigated with an ИСП—51 glass, three-prism spectrograph. The various elements were assembled as shown in Fig. 1, wherein:

Q — denotes the light source,
D — the diaphragm limiting part of the light flow,
F — Schott and chlorine filters,
L_1 — a convergent quartz lenz,
H — a container for holding the luminophor,
L_2 — a convergent glass lens,
S — the spectrograph slit.

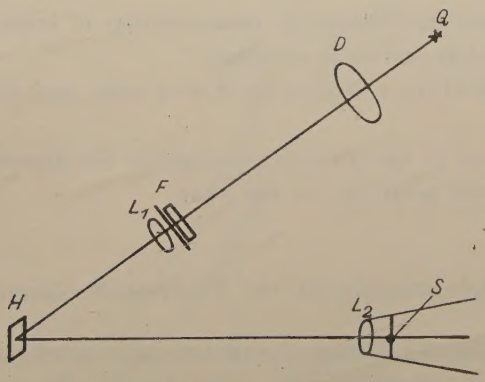

Fig. 1. Scheme of apparatus used in investigation.

The luminophor was illuminated laterally with a TUV—30 Phillips bactericide lamp. The exciting light vonsisted of the mercury line of 2537 Å wavelength, the

[1]) The true content of chlorine and the activators in the luminophors is at present the obiect of investigation.

remaining lines being absorbed in the Schott and chlorine filters. The spectrograph slit width amounted to 0.1 mm. The time of illumination ranged from 2 to 20 minutes. The lines obtained from a helium filled tube served for reference. Polish made "Ultrapan Super" plates were used for the photographs. The luminophors were investigated at the constant temperature of — 20°C.

3. Determination of the Intensity Distribution

To this aim, two series of photographs were taken, one with a graduated attenuator for transverse photometrisation, and one without the attenuator, for longitudinal photometrisation. On plotting the curves of characteristic blakening, the intensity distribution was determined for the wavelengths along which transverse photometry of the plates had been carried out. Longitudinal photometrisation was used for interpolation. On one and the same plate were photographed: the lines of reference, the spectrum of the normal lamp (which was the control spectrum), and the fluorescence spectrum of the respective group of luminophors. The plates were photometrized on a Kipp automatic microphotometer. The relative distribution of the intensiteis of the normal lamp spectrum was determined from Wien's formula (Ornstein 1958)

$$I(\lambda) = C_1 \lambda^{-5} \exp\left(-\frac{C_2}{\lambda T}\right) \tag{3.1}$$

wherein $I(\lambda)$ is the radiation intensity for a given wavelength λ, and C_1, C_2 are constants;
C_1 does not intervene in comparative measurements of intesity ratios,
$C_2 = 1.438$ cm. grad (a universal constant),
T is the temperature of the colour in the Kelvin scale, and was determined as $T = (2840 \pm 10)$°K.

The method used in the present investigation for determining the intensities eliminated the spectral sensitivity of the plate.

4. Determination of the Trichromatic Coordinates

From the intensity distribution, the trichromatic coordinates x, y, z were determined for the various luminophors; this —yielded an objective measure of the colouring of fluorecsence, which plays an important in lighting techniques.

The following formulas were used:

$$x' = \int I_{\text{lum}}(\lambda)\,\bar{x}(\lambda)\,I(\lambda)\,d\lambda; \quad y' = \int I_{\text{lum}}(\lambda)\,\bar{y}(\lambda)\,I(\lambda)\,d\lambda; \quad z' = \int I_{\text{lum}}(\lambda)\,\bar{z}(\lambda)\,I(\lambda)d\lambda \tag{4.1}$$

wherein $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$ are the coordinates of the isoenergetic source, and $I(\lambda)$ — the intensity of source A of colour temperature 2854°K.

x, y, z as functions of x', y', z' are given by the formulas

$$x = \frac{x'}{x' + y' + z'}; \quad y = \frac{y'}{x' + y' + z'}; \quad z = \frac{z'}{x' + y' + z'} \tag{4.2}$$

with

$$x + y + z = 1 \tag{4.3}$$

5. Account of the Fluorescence in the Various Groups of Luminophors

a. Group I.

Group I contains 15 antimony activated luminophors, the amount of Sb increasing from 0 to 7.5%. Luminophor I-1 contains no activator and does not emit light

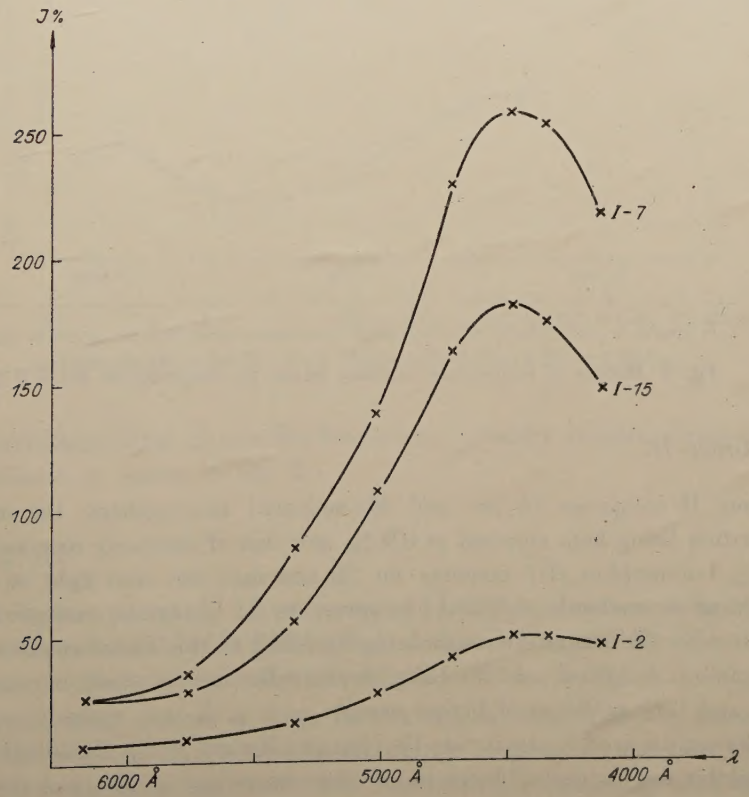

Fig. 2. Example of curves of fluorescence inensity distribution in luminophors of group I, *versus* the wavelength. Mn = 0 %. I—2 Sb = 0.05 %, I—7 Sb = 1.5 %, I—15 Sb = 7.5 %.

on excitation. The distribution of fluorescence of the other luminophors over the various wavelengths is shown in Fig 2.

The graphs show that the luminescence is blue, with maximum intensity at the wavelength of 4472 Å. The dependence of the intensity maxima on the Sb concentration is illustrated in Fig. 3.

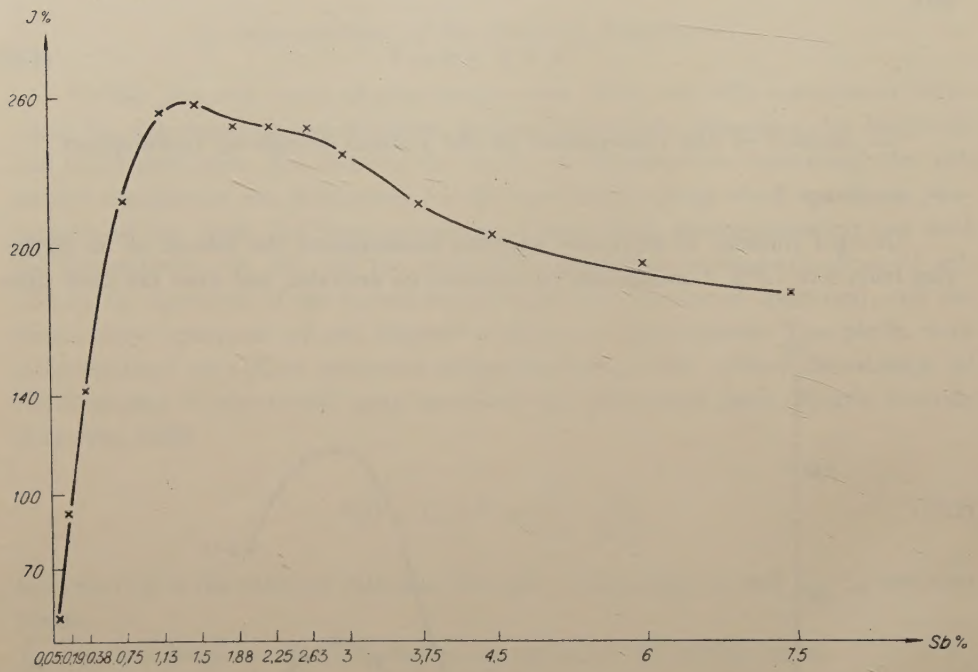

Fig. 3. Maxima of fluorescence intensity *versus* Sb concentration Mn = 0 %.

b. Group II.

Group II comprises 15 Sb- and Mn-activated luminophors, the manganium concentration being kept constant at 0.9 %, and that of antimony increasing from 0 to 7.5 %. Luminophor II-1 contains no Sb and does not emit light on excitation with light of a wavelenth of 2537 Å; however, on for ultraviolet excitation of about 2000 Å, it emits the orange fluorescence characteristic of Mn. The absorption spectrum of manganium activated calcium halophosphate lies in the spectral range beyond 2200 Å, and thus radiation of higher energy, as e. g. X-rays, cathode rays and far ultraviolet can be used for excitation (Botden and Kröger 1948). Luminophors of this group exhibit two ranges of fluorescence: the orange one of Mn, and the blue due to Sb. About 5340 Å, the emission band of the two activators overlap. The intensity versus the wavelength is given in Fig. 4.

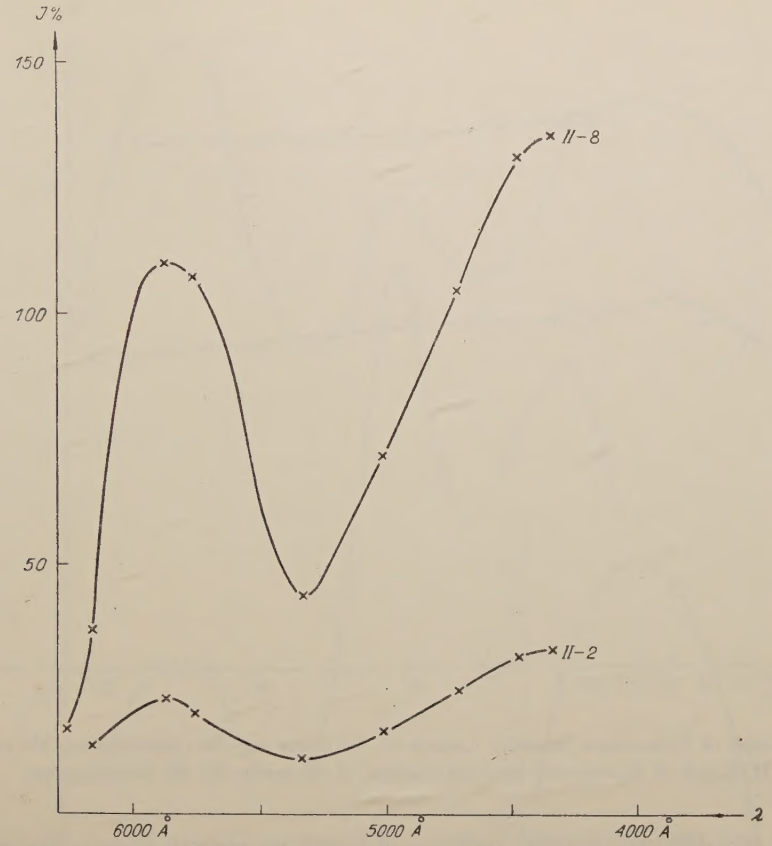

Fig. 4. Example of curves of fluorescence intensity distribution in luminophors of group II, *versus* the wavelength. Mn = 0.9 %. II—2 Sb = 0.05 %. II—8 Sb = 1.88 %.

The distribution of the Sb and Mn fluorescence intensity maxima, versus the rate of Sb introduced, is shown in Fig. 5.

c. Group III.

The third group comprises 13 antimony- and manganium-activated luminophors containing the constant amount of 1.88 % Sb, and wherein Mn varied from 0 to 4.05 %. The intensity distribution versus the wavelength is shown in Fig. 6.

The fluorescence consists of three bands: one deriving from the manganium and having its maximum in the same spectral range as group II, and two belonging to the antimony exhibiting maxima at 5016 Å and 4336 Å. The manganium emission intensity increases steeply with the amount of this activator, and attains its maximum value in luminophor III-8 at 1.8 % Mn, whereafter it decreases slowly. The antimony

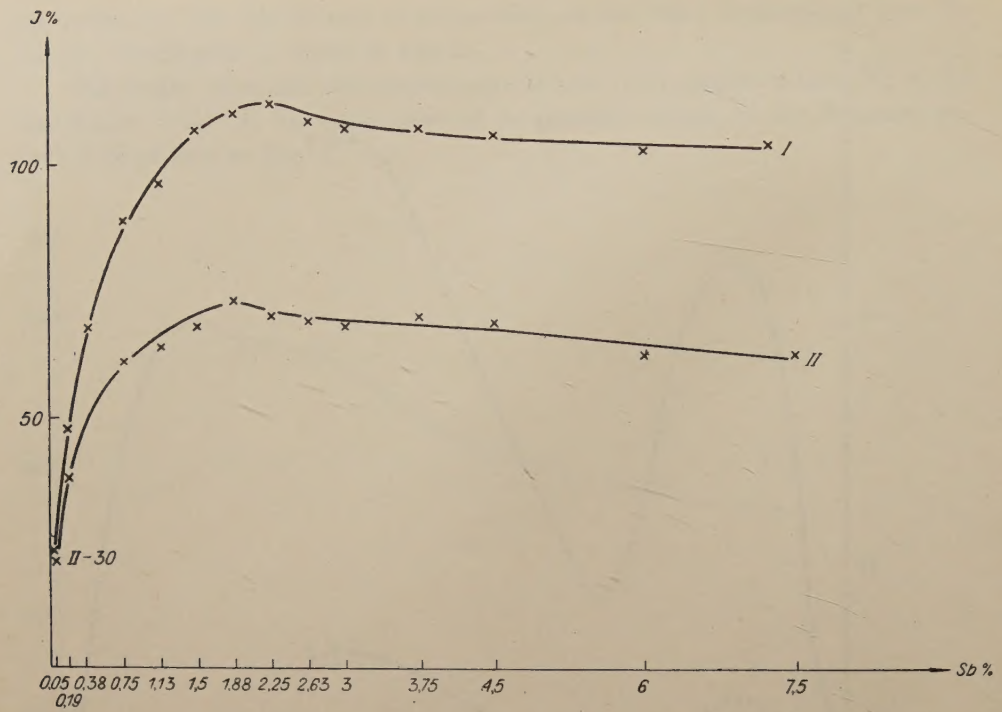

Fig. 5. I Graph of fluorescence intensity maxima of Sb *versus* the Sb concentration. Mn = 0.9 %. II. Graph of fluorescence intensity maxima of Sb *versus* the Sb concentration.

bands are of a different charcter: radiation of longer wavelength appears first in luminophor III-3 and persists as far as luminophor III-10. Its intensity decreases rapidly. The band of shorter wavelength appears in all luminophors. The decrease of the fluorescence intensity occurs exponentially. The distribution of the intensity maxima, as dependent on the amount of Mn incorporated, is shown in Fig. 7a and 7b.

d. Group V.

The fifth group contains five luminophors of the constant activator contents of: Sb = 2.25 %, Mn = 0.9 %, and of variable fluorine to chlorine weight ratio.

The intensity distribution versus the wavelength is shown in Fig. 8.

The distribution of the intensity maxima of the Mn and Sb emission, as versus the weight of the halogens, is shown in Fig. 9a and 9b.

From Fig. 8, a displacement of the Mn fluorescence maxima in the range of 5764—5935 Å is clearly seen to occur when fluorine is entirely replaced by chlorine. However, this causes the maximum of Sb emission to displace but insignificantly towards longer wavelengths. The strongest intensity of Mn and Sb fluorescence occurs in luminophor V-2 of fluorine to chlorine weight ratio 2 : 1.

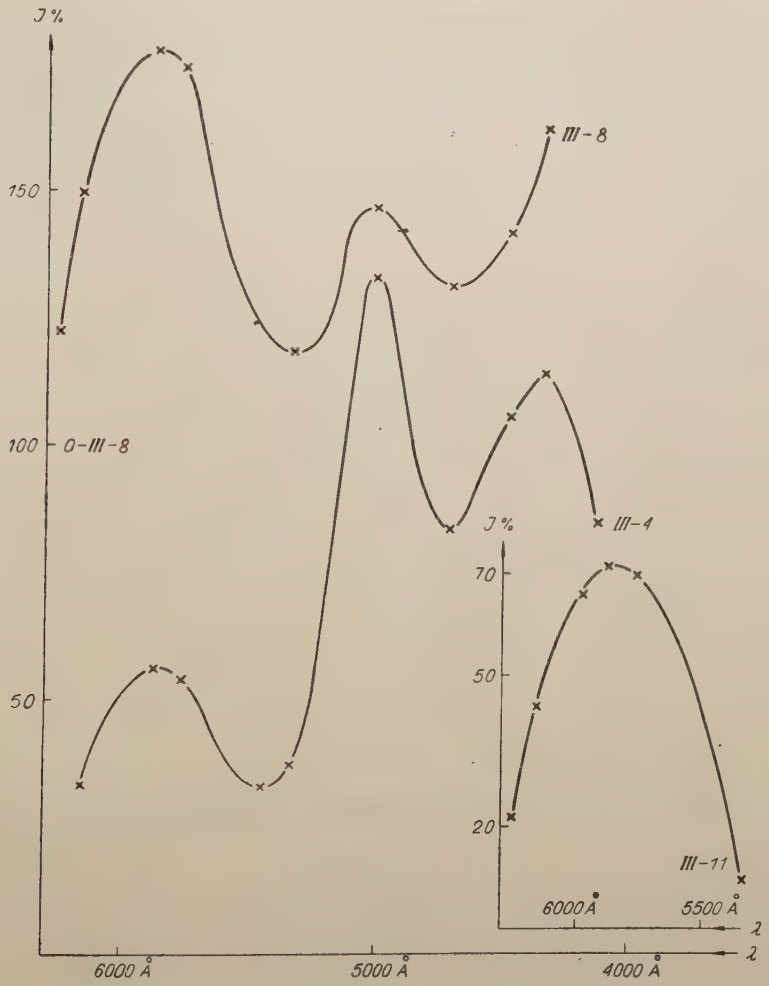

Fig. 6. Example of curves of fluorescence intensity distribution in luminophors of group III, *versus* the wavelength.

III—4 Sb = 1.88 % Mn = 0.6 %, III—8 Sb = 1.88 % Mn = 1.8 %, III—11 Sb = 1.88 % Mn = 3 %.

6. Theoretical Interpretation of Luminescence Mechanism in Luminophors Investigated

A crystal wherein transition of an atom or ion into an excited state has occurred by absorption of radiation or otherwise, can return to the normal state by ridding itself of the energy absorbed. The latter usually is dissipated in the form of heat. In some cases, however, it is re-emitted in that of radiation, and the substance exhibits luminescent properties (Mott 1956). It is seldom that these are exhibited by pure

substances. This is the case of manganium halogenides, rare earths salts, platinocyanides, wolframates, molybdates, and uranyl salts. Randall (1939) suggested that a pure substance is fluorescent if each elementary cell of the crystal contains an ion

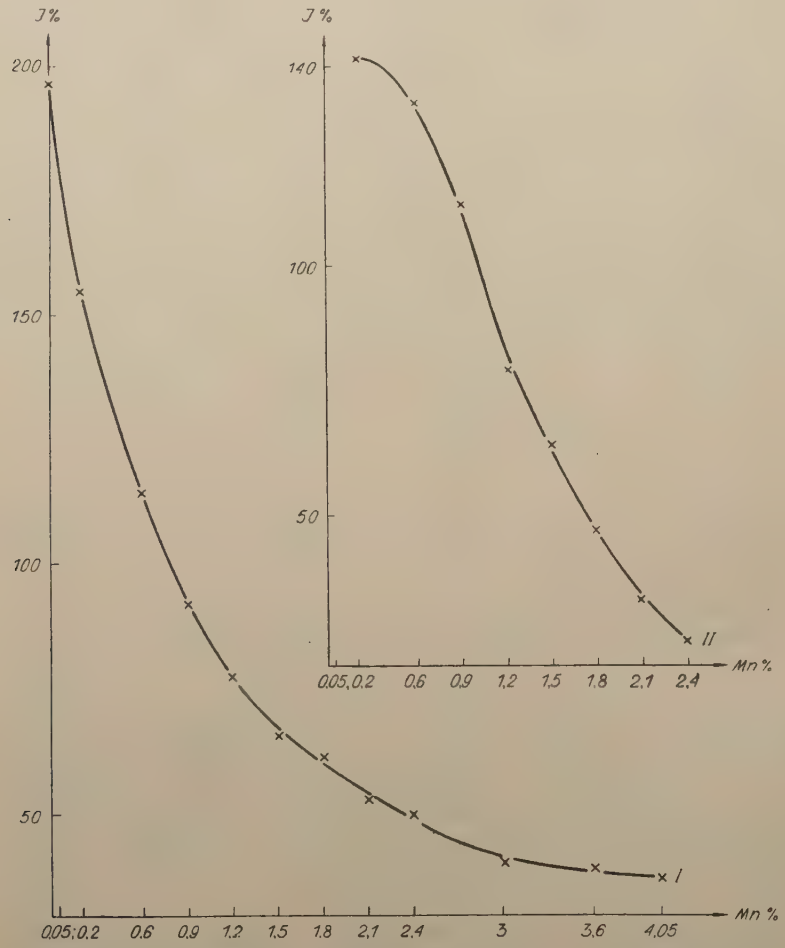

Fig. 7a. I. Fluorescence intensity maxima of Sb, as *versus* Mn concentration, for light of 4336 Å. Sb = = 1.88 %. II. Fluorescence intensity maxima of Sb, as *versus* concentration of Mn, for light of 5016 Å. Sb = 1.88 %.

or complex group with a not fully invested electron layer entirely shielded from the surrounding medium.

From a theoretical point of view, absorption of a radiation quantum by a pure crystal gives rise to a free electron and a hole which recombine without radiation emission, as they are subject to the effect of the crystal lattice vibrations and transfer their energy to the lattice in the form of heat. Hence it results that the electron and

hole should be shielded off from the lattice vibrations. It is here that the crystal lattice defects intervene. At the usual temperatures no ideal lattices exist, since defects occur therein as the vibrations can cause an atom or ion of the crystal to depart from its usual lattice point and to migrate to an interstitial site or to the crystal surface. In certain circumstances, foreign atoms or ions, termed activators, can be incorporated at the void crystal points, forming centres wherein the electron and hole recombine

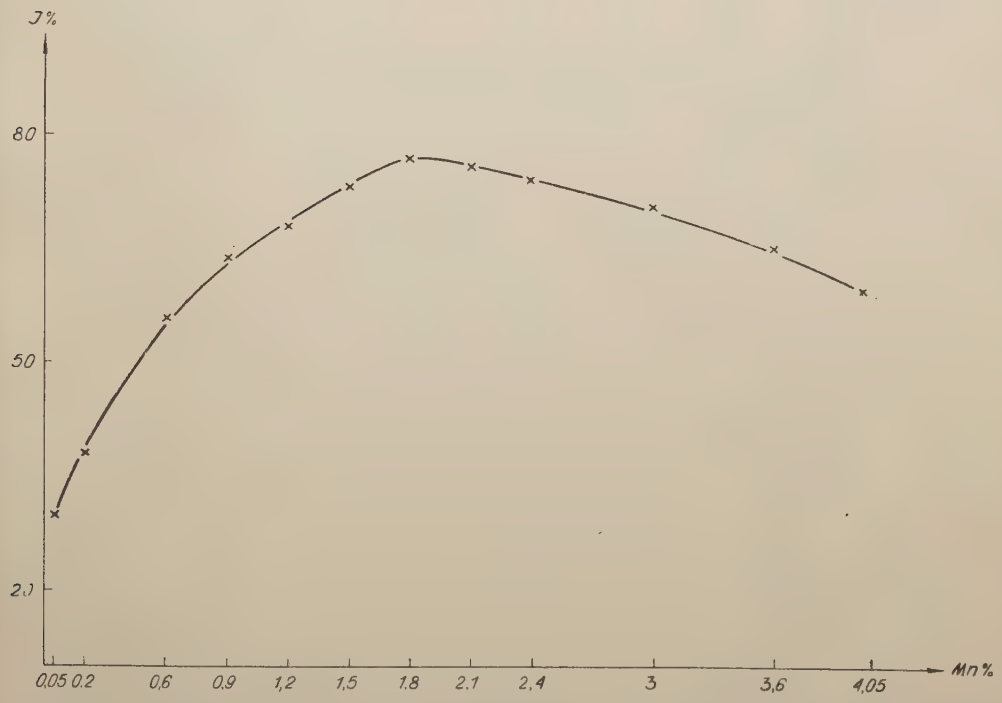

Fig. 7b. Fluorescence intensity maxima of Mn, as *versus* concentration of Mn. Sb = 1.88 %.

with the emission of radiation. The reason for this is that, in a manner hitherto not fully elucidated, the electron and hole are better shielded off at the conters than elsewhere from the effect of the lattice vibrations and thus the probability for them to transfer their energy to the surrounding atoms or ions in the from of heat is lesser.

The conductivity band is essential for the mechanism of luminescence of luminophors having a single activator. Two cases can occur here (Jabłoński 1950):

1. Absorption of exciting radiation is related to the transition of an electron from the activator level to that of conductivity. Emission occurs when one of the electrons of the conductivity level (an excited electron) falls to onto the activator level (Fig. 10).

2. Excitation occurs by the transfer of an electron from the invested fundamental band to the conductivity band. An electron from the invested activator level occupies

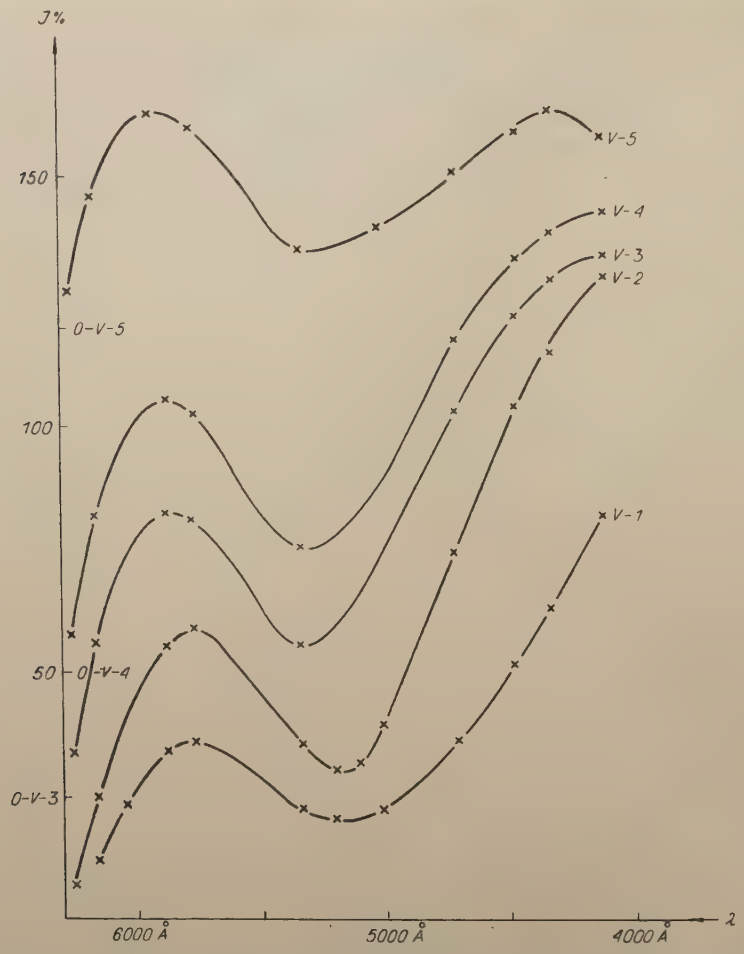

Fig. 8. Curves of fluorescence intensity distribution in luminophors of group V *versus* the wavelength
$F_2 \rightarrow 100-0$ %. $Cl_2 \rightarrow 0-100$ %.
Sb = 2.25 %. Mn = 0.9 %.

a free place in the fundamental band. Emission occurs by the fall of one of the electrons from the conductivity band onto the activator level freed (Fig. 11).

In either case, fluorescence is accompanied by photoconductivity.

N — invested band ,
S — conductivity band,
A — activator level.

Let us now consider the mechanism of light emission in calcium halophosphates. The latter present apatite structure of the ionic type and are doubly activated luminophors. The activators are antimony and manganium. Antimony is termed the

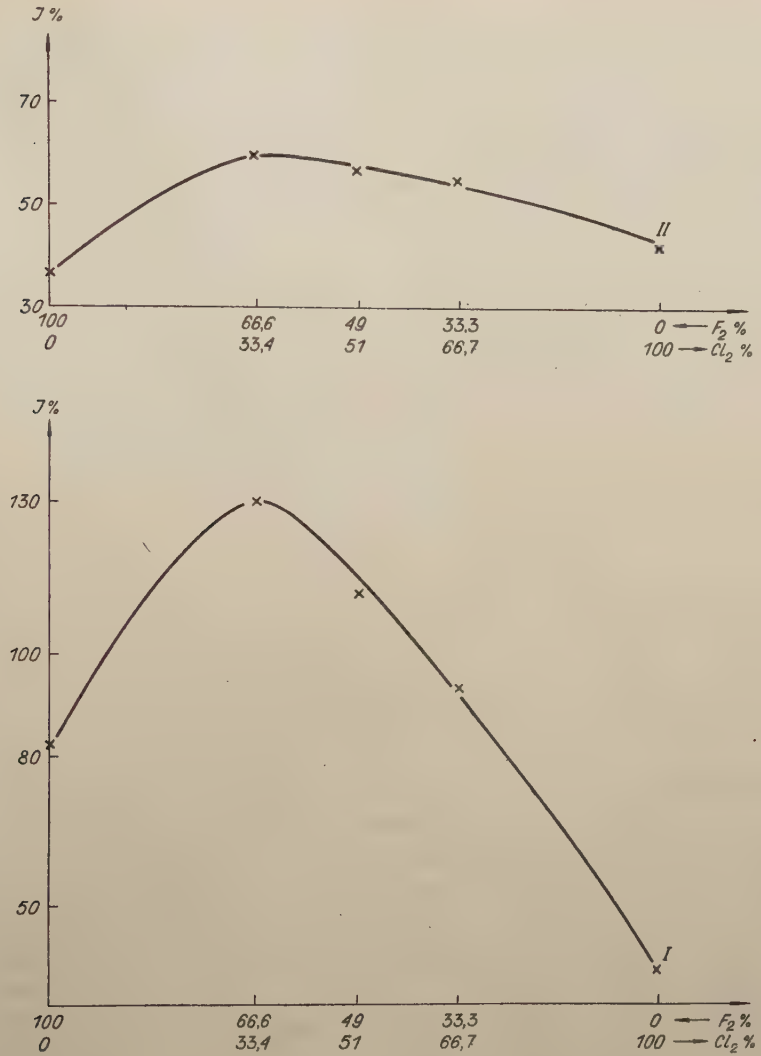

Fig. 9. II. Fluorescence intensity maxima of Sb, as *versus* the concentrations of fluorine and chlorine. II. Fluorescence intensity maxima of Mn, as *versus* the concentrations of fluorine and chorine.

primary activator or sensitizer, and manganium the secondary activator or activator proper. Such luminophors emit what is known as sensitized fluorescence. According to the general theory of sensitized luminescence, the activators replace cations of the basic substances (Williams 1954). Thus, the ions of the two activators replace calcium cations. Manganium and antimony are incorporated in the lattice in the form of bivalent and trivalent ions, respectively. Together with the antimony ion, an oxygen ion is incorporated replacing a halogen ion.

In the present investigation, the luminophors were excited with the mercury 2537 Å line. The role of antimony as sensitizer in this group of luminophors consists in bringing the absorption spectrum into the range of excitation, since neither the basic substance by itself nor when manganium-activated absorbs within this range. Part of the energy absorbed by the sensitizer is emitted in the form of its specific fluorescence, while part is transferred to the activator and emitted in that of the emission spectrum characteristic of the latter, i. e. of manganium. The fluorescence of the

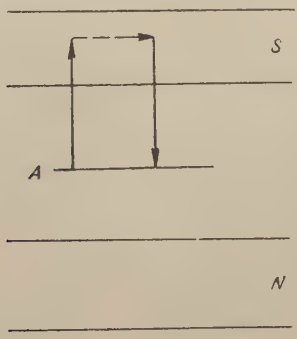

Fig. 10

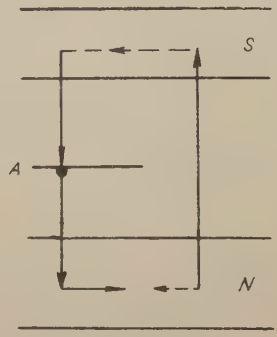

Fig. 11

Fig. 10. Diagram showing fluorescence mechanism in electron transition from activator level to conductivity band.

Fig. 11. Diagram showing fluorescence mechanism in electron transition from fully invested band to conductivity band.

luminophor is always complex, and consists of the emission bands of either activator, the ratio of the intensities depending on the concentrations of the latter.

A problem of far reaching significance arises at this stage, namely, that of the transfer mechanism of energy from the sensitizer to the activator. The literature brings a number of theories aiming at an explanation of the phenomenon of sensitisation. Some had to be rejected as being at variance with the experimental facts. Three examples will be given.

1. The hypothesis of energy transfer from sensitizer to activator through the mediation of the conductivity band had to be rejected, since experiments proved that sensitized luminophors do not exhibit photoconductivity.

2. The hypothesis of energy transfer from excited sensitizer ion to activator ion over the ground state or excited state of the crystal lattice of the principal substance had to be rejected, too, as requiring a quadratic relation between the quantum yield of fluorescence and the intensity of the exciting radiation.

3. The hypothesis of the formation of complex sensitizer and activator centres, termed coactivators, is also at variance with the experimental results (Botden 1952).

The most plansible theory consists in direct energy transfer from sensitizer to activator in a resonance process resembling the one occurring in liquids and gases. We now introduce the following two assumptions:

1. A random distribution of ions of the sensitizer and activator throughout the luminophor is assumed, and

2. Resonance energy transfer between sensitizer-sensitizer and sensitizer-activator ions is assumed (1952a). Energy is transfered from an excited ion of the sensitizer to an activator ion not only by direct resonace involving the two ions, but also over other ions of the sensitizer. Hence, energy absorbed by the sensitizer at a considerable distance from the activator can be transferred to the latter by other sensitizer ions ("enrgy transfer").

We shall now proceed to derive a formula for the energy transfer efficiency. The latter is defined as the ratio of the quantum fluorescence yield of the activator and the sum of those of the sensitizer and activator:

$$Z = \frac{A}{A + S} \tag{6.1}$$

wherein Z denotes the energy transfer officiency,

S — the quantum fluorescence yield of the sensitizer,

A — the quantum fluorescence yield of the activator.

Let us consider all possible paths of energy transfer from sensitizer to activator, namely:

a) an S ion that can transfer energy directly to A (the $S_0 \rightarrow A$ process).

b) an S ion whence energy can be transferred to A exclusively through the intervention of another S ion (the $S_0 \rightarrow S_1 \rightarrow A$ process).

c) an S ion from which energy transfer to A can occur only over two S ions (the $S_0 \rightarrow S_1 \rightarrow S_2 \rightarrow A$ process).

Let k denote the number of cation loci surrounding S and having the property that energy can be transferred from S to A if A occupies one of the k-th loci.

The probability for energy transfer is given by the circumstance that whereas absorption of excitation energy at S situated at any one of the k loci from A involves fluorescence of A, absorption at an S further remote can lead to fluorescence of S.

Let X_a and X_s denote the ratio of the number of ions of the activator and sensitizer, respectively, to the number of cation loci in the crystal. By assumption (1), the probability for finding A at a given distance from S is X_a, whereas the probability that A is not situated at this particular point is $1-X_a$. Then $P_1 = (1-X_a)^k$ means that at each of the k loci not A but some other ion is situated, and $P_2 = 1-(1-X_a)^k$ is the probability for finding A at the k-th locus.

The contribution to Z from the $S_0 \rightarrow A$ process described in (a) is given by the ratio P_2: $(P_1 + P_2)$, i. e. by

$$1-(1-X_a)^k \tag{6.2}$$

Let l denote the number of cation loci surrounding S and having the property of energy transfer between one S and another S situated at one of the l-th loci, and

p — the probability of such transfer. Then the contribution to Z from the $S_0 \to S_1 \to A$ process of (b) is given by the expression.

$$\underbrace{(1-X_a)^k}_{\text{I}}\, p\, \underbrace{[1-(1-X_s)^l]}_{\text{II}}\, \underbrace{[1-(1-X_a)^{k'}]}_{\text{III}} \tag{6.3}$$

wherein I is the probability that there is no A at any of the k-th loci surrounding the initially excited S_0, II is the probability for S_1 to be present at one of the k-th loci around S_0 and for energy transfer from S_0 to S_1, and III — the probability that A is to be found at the k'-th locus in the neighbourhood of S_1, and that the energy received by S_1 is transferred to A. Expression III is analogous to that of (6.2) wherein k has been replaced by k', since some of the k' loci around S_1 can be among the k loci surrounding S_0; the number of overlapping loci is $(k-k')$.

Similarly, the contribution to Z of the $S_0 \to S_1 \to S_2 \to A$ process (c) is given by

$$\underbrace{(1-X_a)^k p\,[1-(1-X_s)^l]\,(1-X_a)^{k'}}_{\text{I}}\, \underbrace{p\,[1-(1-X_s)^l]}_{\text{II}}\, \underbrace{[1-(1-X_a)^{k''}]}_{\text{III}} \tag{6.4}$$

wherein I is the probability for $S_0 \to S_1$ energy transfer from a given locus l in the neighbourhood of S_0 and for no A to be present at any of the loci k or k' around S_0 or S_1, II is that for the presence of an S_2 at one of the l-th loci around S_1 and for the occurrence of $S_1 \to S_2$ energy transfer, and III — the probability for an A to be situated at one of the k''-th loci around S_2, accounting for overlapping of loci of this type and those of types k and k' around S_0 and S_1.

Similarly, the contribution to Z from the $S_0 \to S_1 \to S_2 \to S_3 \to A$ process etc., can be computed.

Adding expressions (6.2), (6.3) and (6.4), we have

$$kX_a - \frac{k(k-1)}{2} X_a^2 + plk' X_a X_s + \frac{k(k-1)(k-2)}{3!} X_a^3 - X_a^2 X_s plk' \left(\frac{k'-1}{2} + k\right) -$$
$$- X_a X_s pl \left[\frac{k'(l-1)}{2} - plk''\right] + \ldots \tag{6.5}$$

If X_a is small with respect to $1/k$ and X_s is small with respect to $1/l$, the terms in (6.4) decrease with higher powers of X_a and X_s.

Denoting by Z_1, Z_2 and Z_3 consecutive approximations to Z, the index equalling the maximum value of the sum of the powers of X_a and X_s, we have

$$Z_1 = kX_a \tag{6.6}$$

$$Z_2 = kX_a - \frac{k(k-1)}{2} X_a^2 + plk' X_a X_s. \tag{6.7}$$

$$Z_3 = kX_a - \frac{k(k-1)}{2} X_a^2 + \frac{k(k-1)(k-2)}{3!} X_a^3 +$$
$$+ plk' X_a X_s \left[1 - X_a \left(\frac{k'-1}{2} + k\right) + X_s \left(\frac{plk''}{k'} - \frac{l-1}{2}\right)\right]. \tag{6.8}$$

By eq. (6.6), the quantity Z does not depend on the concentration X_s of the sensitizer, and is constant for a given concentration X_a of the activator. Hence,

$$k = \frac{Z}{X_a} \tag{6.9}$$

This state can be attained if the concentration of the sensitizer is high as compared to that of the activator and all the activator ions lie within the reach of the sensitizer. It is then that direct $S_0 \rightarrow A$ energy transfer takes place. If this can be attained experimentally, thequantity k can be computed from eq. (6.9).

By eq. (6.7), Z is seen to increase linearly with X_s if X_a is kept at a constant value. Substituting $X_s = 0$ in eq. (6.7), we have

$$Z_{(X_s=0)} = kX_a - \frac{k(k-1)}{2} X_a^2$$

whence

$$k = \frac{1 - [(1-2Z)_{X_s=0}]^{1/2}}{X_a} \tag{6.10}$$

From eq. (6.10), k can be computed. In practice this is possible if the concentration of the sensitizer is small as compared with that of the activator, and not all the ions of the activator are within the range of action of the sensitizer ion. Then energy transfer takes place through the agency of another sensitizer ion, according to the process $S_0 \rightarrow S_1 \rightarrow A$.

By eq. (6.8), Z is a quadratic function of X_s. This corresponds to a process wherein energy can be transferred from S to A exclusively over two S ions, i. e. according to $S_0 \rightarrow S_1 \rightarrow S_2 \rightarrow A$. This should be expected to occur at the very lowest concentrations of the sensitizer.

The foreging theoretical considerations will now be applied to calcium halophosphate luminophors.

By simple considerations, the efficiency of energy transfer, Z, equals the ratio of the fluorescence intensity of the activator and the sum of those of the activator, and sensitizer:

$$Z = \frac{I_{Mn}}{I_{Mn} + I_{Sb}} \tag{6.11}$$

with I_{Mn} and I_{Sb} denoting the intensities of emission of Mn and Sb, respectively.

It result that, if the distribution curves of the intensities are known the quantity Z can be computed. The numerical values, as dependent on the Sb concentration, are shown in Fig. 12.

From Fig. 12 the curve of efficiency of energy transfer, Z, as dependent on the concentration of antimony, si seen to consist of three sections: section I, which corresponds to the lowest Sb concentrations, is an ascending curve and accounts for

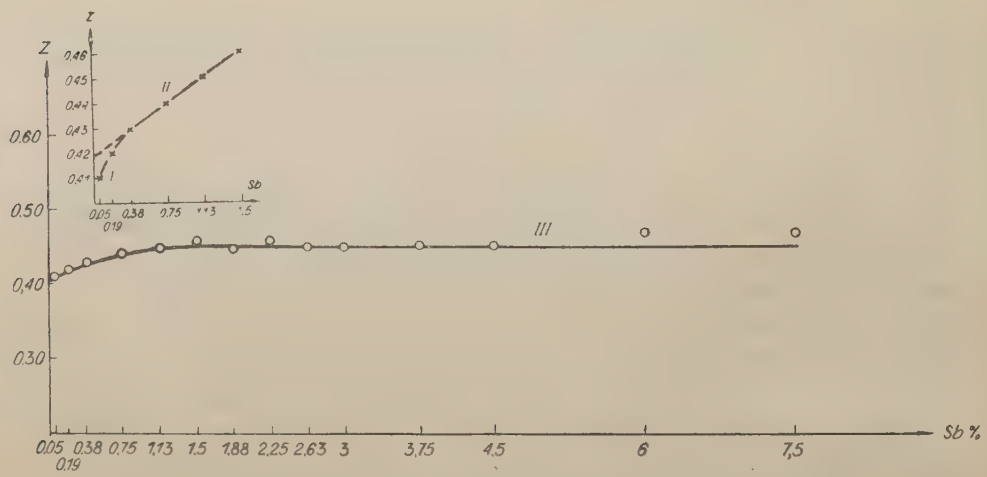

Fig. 12. Energy transfer efficiency *versus* Sb concentration. Mn = 0.9 %.

the $S_0 \rightarrow S_1 \rightarrow S_2 \rightarrow A$ process of energy transfer; section II, for Sb concentrations ranging from 0.38 to 1.13 %, is an ascending straight line accounting for the $S_0 \rightarrow S_1 \rightarrow A$ process; section III, for Sb concentrations of from 1.13 to 7.5 %, is a straight line parallel to the Sb concentration axis and corresponds to the $S_0 \rightarrow A$ process of energy transfer. Thus it is to be concluded that the foregoing theoretical considerations are in agreement with the experimental results.

The value of k as computed by eq. (6.10) from section II of the curve amounts to 61, whereas that computed by eq. (6.9) from section III is 50. The latter is the move reliable of the values, seeing that former is obtained by interpolation of the curve down to $X_s = 0$ and results from no more than 4 points. Botden (1952) obtained values of k for Mn ranging from 26 to 57, as computed from eq. (6.10), with several sensitized systems. However, Botden had insufficient experimental material at his disposal, and in none of the systems investigated did he attain independence of the efficiency of energy transfer with respect to the sensitizer concentration according to eq. (6.6).

Dexter (1953), from theoretical computations of dipole radiation, obtained a value of amounting to 30 for Mn.

Similar considerations hold for the III group of the luminophors containing a constant concentration of Sb and variable Mn.

By eq. (6.6), Z should now be expected to increase linearly with the activator concentration. Results as computed for Z versus X_a are shown in Fig. 13.

The values of the efficiency of energy transfer as computed for the III group of luminophors corroborates the foregoing conclusions. For the Sb concentration, which is 1,88 %, energy transfer is seen to occur according to the equation $Z = kX_a$, i. e. the $S_0 \rightarrow A$ process. For Mn concentrations exceeding 3 %, transfer efficiency

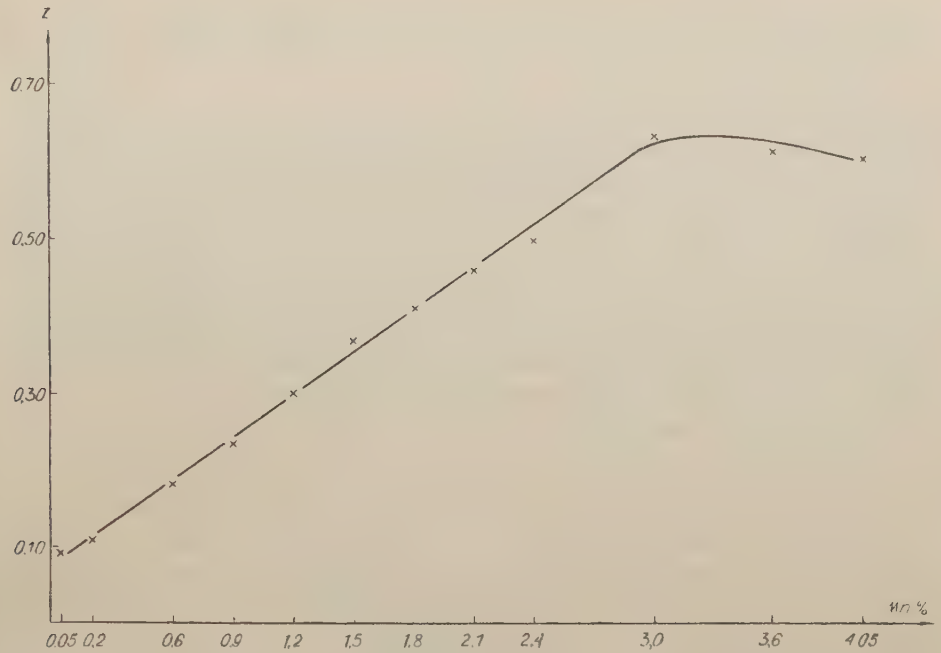

Fig. 13. Energy transfer efficiency *versus* Mn concentration. Sb = 1.88 %.

decreases. Presumably what is known as fluorescence quenching due to activator concentration occurs heve, i. e. the effective number of emitting centres diminishes. According to Schulman (1954), the activator can absorb but cannot emit when within the range of action of another activator ion. Since the fraction of activator ions within the range of another activator ion increases with the concentration, the fluorescence yield must necessarily decrease towards higher concentrations.

7. *Theory of Resonance Energy Transfer in Solids*

The significance of sensitized luminescence in inorganic solids resides in the fact that resonance energy transfer can take place between a permitted transition in the sensitizer and a forbidden transition in the activator. Thus the activator in the present instance manganium yields luminescence subsequent to the transfer of the energy absorbed by the sensitizer, since, by itself, it does not absorb directly. According to Dexter (1953), the entire process of resonant transfer of energy comprises the following five stages:

1. absorption of a photon of energy E_0 by the sensitizer,

2. relaxation of the crystal lattice around the sensitizer ion, to such an extent that the available electron energy in radiative transfer from the sensitizer shall be $E_1 < E_0$,

3. transfer of the energy E_1 to the activator,

4. relaxation of the crystal lattice around the activator ion to such an extent that the available energy emitted in radiation shall be $E_2 < E_1$,

5. emission of energy E_2.

Stages 3 and 4 require 10^{-13} sec each, or maybe longer, for excess energy to be dissipated by the phonons over the crystal lattice, leading to Stokes' law. Stage 3 takes a time that depends on the distance of the surrounding activator ions, i. e. on the concentration X_a. Stage 5 requires 10^{-8} sec or much more, according to the degree of forbiddenness of the transition. Since Stokes' law is usually fulfilled, the activator represents a kind of "trap" for the electron excitation energy, i. e. the excited activator cannot enter into resonance with the sensitizer ($E_2 < E_0$) and no energy transfer can take place from activator to sensitizer. The foregoing considerations serve to explain two facts of the domain of sensitized luminescence, namely:

1. in all sensitized systems, fluorescence of the activator exhibits a longer wavelength than that of the sensitizer;

2. the fluorescence decay time of the sensitizer is much shorter than that of the activator, e. g. amounting to 10^{-6} sec for Sb and to 10^{-2} sec for Mn.

We shall now consider resonance energy transfer from a theoretical point of view, starting from the perturbation theory (Herzberg 1939) of quantum mechanics. By this theory the magnitude of the pertubation between two states depends on the matrix element W_{ni} and on the differences in energy between the non-perturbed levels. The matrix element W_{ni} can divided in two parts: one dependent on the electron and rotatory parts of the wave function, and one dependent on the oscillatory part thereof. The latter part is

$$W_{ni} = \int \psi_n^v \, W^v \, \psi_i^{v*} \, dr \tag{7.1}$$

wherein ψ_n^v and ψ_i^{v*} are oscillatory wave functions in Schrödinger's equation, and W^v is a function of the distances between the nuclei. By eq. (7.1), a strong perturbation occurs only when the oscillatory parts of the wave functions overlap appropriately. This can occur only for two oscillatory levels situated in the neighbourhood of the point of intersection of the potential curves. In luminescennce, a perturbation between two oscillatory states appertaining to two different ions (or atoms) leads to energy transfer from one of the ions (or atoms) to the other. In the perturbating state appertains to an ion of the impurity, usual luminescence quenching occurs. In sensitized luminescence, an ion of the activator proper is the perturbating factor. At the moment of excitation it remains in the normal state, as transition to an excited state is forbidden by the selection rules. However, if it possesses excited states of energy almost equal to that of the excited states of the sensitizer, the energy of the latter is transferred to the activator, raising it to an excited state. Hence, in sensitized luminescence, the roles of sensitizer and activator proper can be played only by ions fulfilling the two conditions:

1. they must exhibit excited states of almost equal energy,

2. for obtaining maximum efficiency of energy transfer, the differences between the various excited energy states, both in the sensitizer and activator, should be small.

Numerous tests proved that these are best fulfilled by the combination of Sb as sensitizer and Mn as activator proper. Antimony can be replaced by arsenic, bismuth, lead or tin, thoung the result is so satisfactory. The only activator proper known in sensitized luminescence is, however, manganium.

Applying the perturbation theory outlined above to sensitized luminescence in solids, the following conclusions may be deduced:

1. Two oscillatory states of sensitizer and activator, of almost equal energy, interact in such a way that energy from the sensitizer is transferred to the activator.

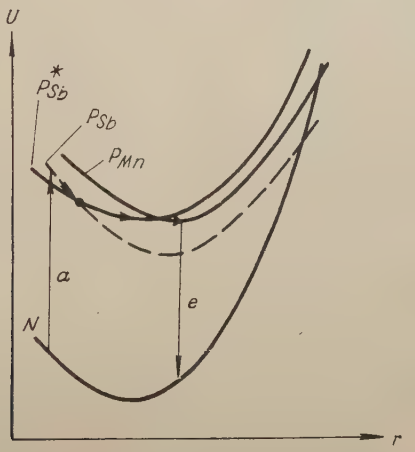

Fig. 14. Curves of potentional energy of fundamental and excited states for antimony and manganium.

2. Maximum interaction occurs if the potential energy curves of either state approach one another or intersect.

3. Hence results the relation between the potential energy curves in the fundamental and excited states for Sb and Mn shown in Fig. 14.

U — potential energy,
r — distance between nuclei,
N — fundamental state,
P_{Mn} — excited state in Mn,
P_{Sb} — excited state in more remote Sb,
P^*_{Sb} — excited state in nearer Sb.

Energy absorbed by the Sb ion very remote from the Mn ion is transferred to a nearer ion, whence it is, in turn, transferred to the Mn ion.

The author wishes to express her very great indebtedness to the Promoter, Professor Dr J. [illegible], to Professor Dr W. [illegible] for giving his permission for carrying out the measurements with the [illegible] at his Institute, and to Professors Dr Dr S. [illegible] and W. [illegible] for their [illegible] remarks in giving their opinions on this paper.

[illegible]

Botden, *Philips Res. Rep.*, **7**, 197 (1952).

Botden and Kröger, *Physica*, **14**, 553 (1948).

Bundel et al., *Izv. Akad. Nauk, SSSR, Ser. fiz.*, **15**, [illegible] (195[illegible]).

Butler and Jerome, *J. Electrochem. Soc.*, **97**, 265 (1950).

Dexter, *J. chem. Phys.*, **21**, 836 (1953).

Doherty and Harrison, *Brit. J. appl. Phys.*, Suppl. No. 4 S 11 (1955).

Gurewicz, *Tsvet i yewo izmerenye* (1950).

Herzberg, *Molecular Spectra and Molecular Structure of Diatomic Molecules* (1939).

Jabłoński, *Postępy Fizyki*, **1**, 163 (1950).

Jenkins, McKeag and Ranby, *J. Electrochem. Soc.*, **96**, 1 (1949).

McKeag and Ranby, *J. Electrochem. Soc.* ([illegible]).

Mott N. and Gurney R., *[illegible]*, Warszawa 1956 PWN (Electron Phenomena in Ionic Crystals, 1956)

Orlstein at al., *[illegible]*, Warszawa 1958 PWN (Objective Spectrophotometry [illegible])

Randall, *Trans Faraday Soc.*, **35**, 2 (1939).

Schulman, *Brit. J. appl. Phys.*, nr 4, 64 (1954).

Williams, *Brit. J. appl. Phys.*, nr 4, 97 (1954).

ON THE POSSIBILITY OF DETERMINING THE DENSITY OF DISLOCATIONS BY MEANS OF AN X-RAY SPECTROGRAPH WITH OSCILLATING FILM*)

By Teresa Bedyńska

Institute of Physics, Polish Academy of Sciences, Warsaw

(Received November 17, 1959)

A method of computing the distribution function of mosaic blocks in the crystal is proposed. For this, it is required that the forms of Bragg's reflex as obtained on an X-ray spectrograph with fixed film, and that of the reflex, as obtained on a spectrograph with oscillating film, be known. If the distribution function of mosaic blocks is known, the density of dislocations can be computed.

1. Introduction

The boundaries between blocks in a crystal of mosaic structure are formed by the dislocation lines. If the crystal consist structurally of blocks of edge length t and if the distance between the dislocation lines on the boundary separating the blocks is h, the density of dislocations on the boundary $\varrho_{\text{bound}} = 1/h^2 = 1/t^2$ (Gay, Hirsh, Kelly 1953). If the blocks in the crystal were formed by a system of parallel dislocation lines having one and the same Burgers' vector, the density ϱ_{bound} as computed from Gay, Hirsh and Kelly's formula (1953) would be numerically equal to the number of dislocation lines intersecting a unit surface parallel to its normal. The angle between the blocks in a crystal $\alpha = b/h$ (Burgers 1940). b denotes the modulus of Burgers' vector. Hence, the densitiy at the boundary between the blocks is

$$\varrho_{\text{bound}} = \frac{\alpha}{bt} \tag{1}$$

In addition to dislocations at the boundaries, dislocation lines can appear within crystal blocks. The total dislocation density for a system of parallel dislocation lines is $\varrho = 1/h_1^2$ with h_1 denoting the mean distance between the dislocations. For mosaic crystals whose mosaic blocks are so small and so weakly disoriented that using the

*) The spectrograph with oscillating film is described in the papers by J. Auleytner (1958, 1959 a, b).

micro-beam technique, reflexes from the various blocks can not be obserwed separatly, Gay, Hirsh and Kelly (1953) assume a system of blocks with a single dislocation per one inter-block boundary. With this assumption, the total dislocation density is given by formula:

$$\varrho = \left(\frac{\alpha}{b}\right)^2 \tag{2}$$

In order to find the dislocation density, the mean angle between the blocks should be known. The latter may be determined if the angular distribution function for the mosaic blocks in the crystal is known. This function can be obtained from analysis of the line form yielded by the X-ray spectrograph with oscillating film (Auleytner 1958, 1959 a, b), and comparing it with the one yielded by a spectrograph with immobile film.

2. Principles of Method of Determining Mosaic Blocks Distribution Function

The principle of the method is given in Fig. 1. An ideally parallel and narrow beam of X-rays is incident at point O on the ideally smooth crystal surface. The rotation axis of the crystal is at point O. The crystal rotates about axis O and displaces parallel to the reflecting surface. On the immobile film F, a line of the form given

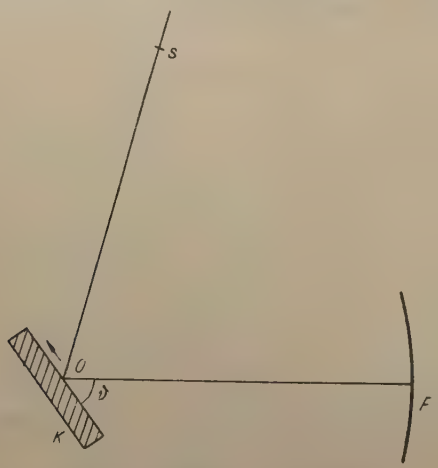

Fig. 1. Principle of the method.

by the function $g(\frac{1}{2}x)$ arises. If the film F rotates simultaneously with the crystal K about O, a line given by the function $h(x)$ results. The functions $g(x)$ and $h(x)$ are related by the following equation:

$$h(x) = \int_{-\infty}^{+\infty} f(y)\, g(x-y)\, dy \tag{3}$$

wherein $f(x)$ is the function of angular distribution of the mosaic blocks.

In experimental investigation, the effects of: 1) divergence of the beam, 2) width of the slit, and 3) inequalities of the crystal surface and displacement of the axis of rotation with respect to the reflecting surface should be accounted for.

3. *Effect of Divergence of Beam in Horizontal Plane on Line Forms Obtained on Immobile and Oscillating Films*

In the case under consideration, a beam, divergent in the horizontal plane (the plane parallel to the rotation axis of the crystal) and parallel to the vertical plane, is incident on the crystal K from an ideally narrow slit S (Fig. 2) on the focussing circumference. The form of the divergent beam is described by the function $\Phi(R\beta)$

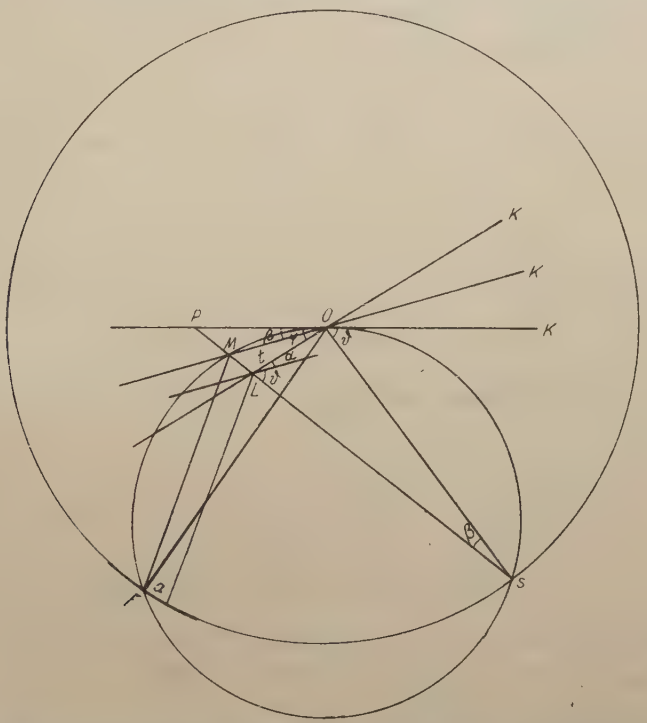

Fig. 2. Effect of divergence of the beam in horizontal plane on forms of the lines.

with β denoting the angle subtended by the beam SP incident on point P of the crystal and the direction of the beam SO incident on its rotation axis, and R denoting the distance of the slit S from the rotation axis O. Let α denote the desorientation angle of the mosaic blocks. The displacement, on the immobile film, of the beam reflected at point L (Fig. 2) with respect to that reflected at O, as resulting from the mosaic, is a:

$$a = 2 \cos \vartheta \, t \sin \alpha$$

wherein ϑ is the angle fulfilling Bragg's equation. and t — the distance of the reflecting point L of the crystal from the axis O:

$$t = \frac{R}{\sin(\vartheta + \alpha)} \sin\beta$$

whence

$$a = 2R \frac{\cos\vartheta}{\sin(\vartheta + \alpha)} \sin\beta \sin\alpha$$

As the crystal K rotates, the angle between its surface and the direction SO varies. Let φ be the angle by which the one subtended by the direction SO and the crystal surface differs from ϑ. The condition for the fulfillment of Bragg's equation for a block deflected by α is given by

$$\alpha = \varphi - \beta \tag{4}$$

The line form, for immobile film, and rotated and displaced crystal, is given by the function $G(x)$:

$$G(x) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g\left[\frac{1}{2}\left(x - 2R\frac{\cos\vartheta}{\sin(\vartheta+\alpha)}\sin\beta\sin(\varphi-\beta)\right)\right] \times$$
$$\times \Phi(R\beta) f(R\varphi - R\beta)\, d(R\varphi)\, d(R\beta) \tag{5}$$

The line form, for mobile film, and rotated and displaced crystal is given by the function $H(x)$:

$$H(x) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g\left[x - 2R\frac{\cos\vartheta}{\sin(\vartheta+\alpha)}\sin\beta\sin(\varphi-\beta) - R\varphi\right] \times$$
$$\times \Phi(R\beta) f(R\varphi - R\beta)\, d(R\varphi)\, d(R\beta) \tag{6}$$

At small values of α and β the term $2R\cos\vartheta\sin\alpha\sin\beta/\sin(\vartheta+\alpha)$ is by orders of magnitude smaller than $R\alpha$ and $R\beta$ and, hence, can be neglected.
Eqs. (5) and (6) assume the form:

$$G(x) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(\tfrac{1}{2}x)\,\Phi(R\beta) f(R\alpha)\, d(R\alpha)\, d(R\beta) = c g(\tfrac{1}{2}x)$$

$$H(x) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x - R\beta - R\alpha)\,\Phi(R\beta) f(R\alpha)\, d(R\alpha)\, d(R\beta)$$

whence

$$H(x) = \int_{-\infty}^{+\infty} h(x - R\beta)\,\Phi(R\beta)\, d(R\beta) \tag{7}$$

wherein

$$h(x) = \int_{-\infty}^{+\infty} g(x - R\alpha)\, f(R\alpha)\, d(R\alpha)$$

The function $h(x)$ is the one that would result on the oscillating film with a parallel beam eq. (3)

4. Effect of Finite Dimensions of Slit on Line Forms $G(x)$ and $H(x)$

If the slit of finite width, the repartition of the intensity thereon is accounted for by the function $\Phi_g(z)$, From each point of the slit S (Fig. 2), a beam diverges. The function $\Phi_z(z+R\beta)$ accounts for the divergence in the horizontal plane of the beam leaving point z of the slit. With such a slit, the line form on the immobile film becomes

$$G(x) = \int_{-\infty}^{+\infty} g\left(\tfrac{1}{2}(x - z)\right) \Phi_g(z)\, dz \tag{8}$$

and, on the oscillating film:

$$H(x) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h(x - R\beta - z)\, \Phi_g(z)\, \Phi_z(z + R\beta)\, dz\, d(R\beta)$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h(x - t)\, \Phi_z(t)\, \Phi_g(z)\, dz\, dt$$

whence

$$H(x) = \int_{-\infty}^{+\infty} h(x - t)\, \Phi_h(t)\, dt \tag{9}$$

wherein

$$\Phi_h(t) = \int_{-\infty}^{+\infty} \Phi_g(z)\, \Phi_z(t)\, dz$$

The function $\Phi_h(t)$ describes the form of the beam leaving the slit, at a distance therefrom equal to that of the crystal from the slit.

5. Effect of Inequalities in Crystal Surface and of Displacement of Rotation Axis with Respect to Crystal Surface

Inequalities inherent in the crystal surface, and displacement of the axis of rotation with respect to the crystal surface can result in some additional widening of the lines $G(x)$ and $H(x)$.

Displacement of the crystal surface by the amount t causes the line $g(x)$ on the film to displace by d (Fig. 3). $g(x)$ describes the line form that would arise on the immobile film with an ideally narrow and parallel beam:

$$d = 2 \cos \vartheta \, t \tag{10}$$

Displacement of the rotation axis of the crystal with respect to its surface by the amount l gives rise to the variation t_0 in the distance of the reflecting surface during rotation of the crystal (Fig. 4). If γ denotes the angular width of the line reflected, then

$$t_0 = l \operatorname{tg} \frac{\gamma}{4} \sin \frac{\gamma}{2} \approx \frac{1}{8} l \gamma^2 \tag{11}$$

Such difluence in $G(x)$ and $H(x)$, arising from this variation in the distance of the crystal surface, can be neglected.

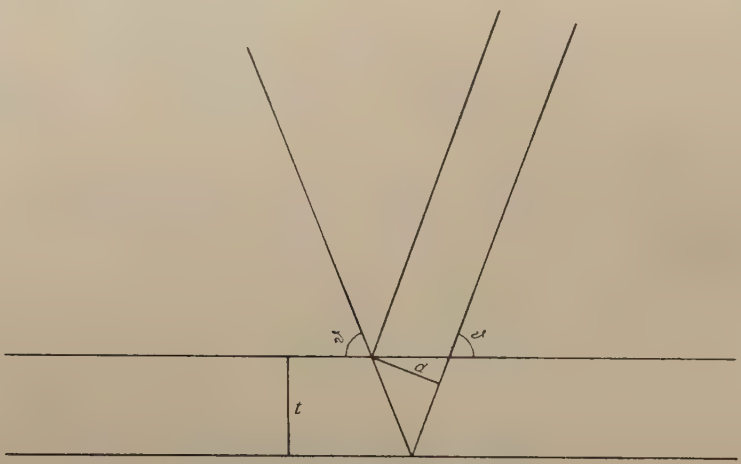

Fig. 3. Effect of inequalities in the crystal surface on forms of the lines.

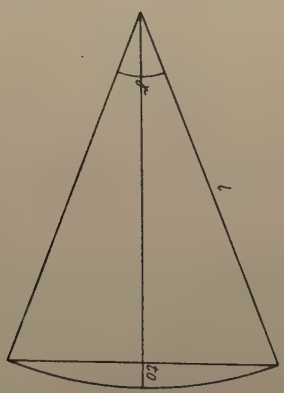

Fig. 4. Effect of displacement of rotation axis with respect to the crystal surface on forms of the lines.

If the widening of the line on the immobile film, as resulting from inequalities in the surface scanned by the X-ray beam when the crystal is not displaced, be neglected, the widening resulting from surface inequalities over a larger area, as investigated when the crystal is subjected to displacement, can be eliminated. In order to eliminate this latter widneing, it is necessary that, in addition to the line forms $H(x)$ obtained with displaced crystal and oscillating film, as well as $G(x)$ obtained with displaced crystal and immobile film, the line form obtained with non-displaced crystal and immobile film should be known. If the surface exhibits inequalities, the line resulting on the immobile film when the crystal is subjected to displacement assumes the form

$$G(x) = \int_{-\infty}^{+\infty} X(t)\, G'(x-t)\, dt \tag{12}$$

The one obtained on the oscillating film as the crystal is subjected to displacement assumes the form

$$H(x) = \int_{-\infty}^{+\infty} X(t)\, H'(x-t)\, dt \tag{13}$$

The functions $G'(x)$ and $H'(x)$ describe the line forms that would arise from the same crystal, assuming its surface to be smooth. Since the inequalities within the limits of the beam, at non-displaced crystal, are neglected here, the function $G'(x)$ will be identified with the function obtained when the crystal is not subjected to displacement. As the functions $G'(x)$, $G(x)$ and $H(x)$ are known experimentally, the unknown function $X(x)$ of the widening arising from surface inequalities can be eliminated. To determine the function of angular distribution of the mosaic blocks, use is made of eqs. (3), (8), (9), (12) and (13):

$$h(x) = \int_{-\infty}^{+\infty} f(y)\, g(x-y)\, dy$$

$$G'(x) = \int_{-\infty}^{+\infty} g\left(\tfrac{1}{2}(x-y)\right) \Phi_g(y)\, dy$$

$$G(x) = \int_{-\infty}^{+\infty} X(y)\, G'(x-y)\, dy$$

$$H'(x) = \int_{-\infty}^{+\infty} h(x-y)\, \Phi_h(y)\, dy$$

$$H(x) = \int_{-\infty}^{+\infty} X(y)\, H'(x-y)\, dy$$

The foregoing system can be replaced by the single equation

$$H^*(x) = \int_{-\infty}^{+\infty} G^*(x-y)\, f(y)\, dy \tag{14}$$

together with the four integrals

$$H^*(x) = \int_{-\infty}^{+\infty} H(x-y)\, F_g(y)\, dy \tag{15}$$

$$G^*(x) = \int_{-\infty}^{+\infty} G(x-y)\, F_h(y)\, dy \tag{16}$$

$$F_g(x) = \int_{-\infty}^{+\infty} G'(x-y)\, \Phi_g(2y)\, dy \tag{17}$$

$$F_h(x) = \int_{-\infty}^{+\infty} G'(2(x-y))\, \Phi_h(y)\, dy \tag{18}$$

$f(y)$ is the distribution function of mosaic blocks required. The following functions are given experimentally: $H(x)$, $G(x)$ $G'(x)$, $\Phi_g(x)$ and $\Phi_h(x)$
$H(x)$ accounts for the line form, at displaced crystal, on oscillating film.
$G(x)$ accounts for the line form, at displaced crystal, on immobile film.
$G'(x)$ accounts for the line form, at non-displaced crystal. on immobile film.
$\Phi_g(x)$ describes the form of the beam leaving the slit, on the latter.
$\Phi_h(x)$ describes the form of beam leaving the slit, at distance therefrom equal to that of the crystal from the slit.

6. *Effect of Divergence within Vertical Plane*

In addition to that arising from divergence within the horizontal plane, considerable modification in the line forms $G(x)$ and $H(x)$ is given rise to by divergence within the vertical plane (the one containing the rotation axis of the crystal). If, the crystal K remains immobile and a beam $\boldsymbol{k}$ substending the angle φ with the beam $\boldsymbol{k}_0$ perpendicular to the slit S and axis O (Fig. 5) in the vertical plane is incident on the crystal, then, in order that both fulfill Bragg's condition simultaneously, $\boldsymbol{k}$ must be deflected by the angle α with respect to $\boldsymbol{k}_0$ in the horizontal plane:

$$\operatorname{tg}\alpha = \frac{\sin\vartheta\cos\vartheta - \sin\vartheta\cos\vartheta\left(1 - \dfrac{\operatorname{tg}^2\varphi}{\cos^2\vartheta}\right)^{1/2}}{1 - \cos^2\vartheta + \cos^2\vartheta\left(1 - \dfrac{\operatorname{tg}^2\varphi}{\cos^2\vartheta}\right)^{1/2}} \tag{19}$$

if $\frac{\operatorname{tg}^2 \varphi}{\cos^2 \vartheta} < 1$

$$\operatorname{tg} \alpha = \frac{\sin \vartheta \cos \vartheta \left(\frac{1}{2} \frac{\operatorname{tg}^2 \varphi}{\cos^2 \vartheta} + \frac{1}{8} \frac{\operatorname{tg}^4 \varphi}{\cos^4 \vartheta} + \frac{1}{16} \frac{\operatorname{tg}^6 \varphi}{\cos^6 \vartheta} + \ldots \right)}{1 - \cos^2 \vartheta \left(\frac{1}{2} \frac{\operatorname{tg}^2 \varphi}{\cos^2 \vartheta} + \frac{1}{8} \frac{\operatorname{tg}^4 \varphi}{\cos^4 \vartheta} + \frac{1}{16} \frac{\operatorname{tg}^6 \varphi}{\cos^6 \vartheta} + \ldots \right)}$$

at small values of φ

$$\operatorname{tg} \alpha \approx \tfrac{1}{2} \operatorname{tg} \vartheta \operatorname{tg}^2 \varphi \, (1 + \tfrac{1}{2} \operatorname{tg}^2 \varphi + \tfrac{1}{4} \operatorname{tg}^4 \varphi + \ldots)$$

$$\alpha \approx \tfrac{1}{2} \operatorname{tg} \vartheta \operatorname{tg}^2 \varphi \tag{20}$$

If the beam exhibits divergence both in the vertical and horizontal planes, a line will result on the photographic plate, of a width determined by the horizontal and vertical divergences of the beam.

Divergence can be removed by reflecting the beam from two immobile crystals.

Using a point focus X-ray tube, it is possible to operate with a beam divergent in the vertical plane and to employ eq. (14) and integrals (15), (16), (17) and (18) in computing the distribution functions of the mosaic blocks. A point focus yields, at each point of the film, an image resulting from the beams exhibiting one and the

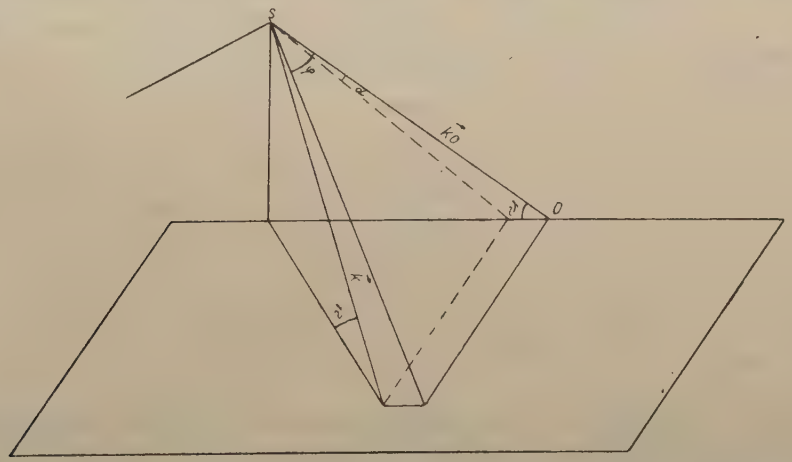

Fig. 5. Effect of divergence of the beam in vertical plane of forms on the lines.

same value of the angle φ. The beam $\boldsymbol{k}$ leaving the focus at an angle φ in the plane perpendicular with respect to the beam normal to the slit and rotation axis of the crystal, fulfills Bragg's conditions if the angle between the crystal surface and $\boldsymbol{k}_0$ amounts to

$$\vartheta - \alpha = \vartheta - \tfrac{1}{2} \operatorname{tg} \vartheta \operatorname{tg}^2 \varphi \tag{21}$$

Variations of the angle $\alpha = \frac{1}{2} \operatorname{tg} \vartheta \operatorname{tg}^2 \varphi$ with ϑ can be neglected within the width of the line investigated. These decrease as the angle φ diminishes. Hence, for any beam $\boldsymbol{k}$ subtending a small angle with $\boldsymbol{k}_0$, the line forms: $G(x)$ resulting on an immobile film, and $H(x)$ — as obtained on an oscillating film — undergo no modification. The lines $G(x)$ and $H(x)$ are displaced. In experimental work, the height of the tube focus and photometer spot should be made so small as to provide for divergence $\Delta\varphi$ in the vertical plane within the height of the photometer slit yielding much smaller displacement of the line than the dispersion of the mosaic blocks investigated. If $\operatorname{tg} \varphi \cos^{-2} \varphi \Delta\varphi$ is considerably smaller than the angle of dispersion of the mosaic blocks, then divergence within the vertical plane will not have contributed towards deforming the lines $G(x)$ and $H(x)$. In this case, the dispersion function can be computed from eqs. (14), (15), (16), (17) and (18).

7. *Effect of Widening from Reflexion at Interior Faces, and from Surface Inequalities*

Eqs. (14), (15), (16), (17) and (18) do not serve the purpose of eliminating widening as arising from the surface inequalities present within the width of the beam, and from reflexion at crystal faces situated at a certain depth beneath the outer surface. This latter circumstance can come to play an important part in crystals exhibiting a low rate of absorption, but can be eliminated by measuring the Bragg line form on an immobile film disposed at one half of the distance from the crystal in mobile film investigation. The block distribution function is now computed from the following formulas:

$$h(x) = \int_{-\infty}^{+\infty} f(y)\, g(x-y)\, dy$$

$$G(x) = \int_{-\infty}^{+\infty} g(x-y)\, \Phi_g(y)\, dy$$

$$H(x) = \int_{-\infty}^{+\infty} h(x-y)\, \Phi_h(y)\, dy$$

With the foregoing system of equations, error arising through surface inequalities and reflexion at interior crystal faces is eliminated. The system of equations can be replaced by the equation

$$H^*(x) = \int_{-\infty}^{+\infty} G^*(x-y)\, f(y)\, dy \tag{22}$$

together with the integrals

$$H^*(x) = \int_{-\infty}^{+\infty} H(y)\, \Phi_g(x-y)\, dy \tag{23}$$

and

$$G^*(x) = \int_{-\infty}^{+\infty} G(y)\, \Phi_h(x-y)\, dy \tag{24}$$

$f(x)$ is the distribution function of the blocks required.
$H(x)$ accounts for the line form on the oscillating film, at distance R from the crystal.
$G(x)$ accounts for the line on the immobile film, at distance $\frac{1}{2} R$ from the crystal.
$\Phi_g(x)$ describes the shape of the beam incident on the crystal, at distance $\frac{1}{2} R$ therefrom.
$\Phi_h(x)$ describes the shape of the beam incident on the crystal, at the latter.

8. On the Form the Block Distribution Function as Resulting from Variously Oriented Systems of Dislocation Lines

The presence of dislocations gives rise to a system of blocks within the crystal. The blocks formed by edge dislocations are mutually rotated about the dislocation line. Those formed by screw dislocations are rotated about the normal to the plane of sliding.

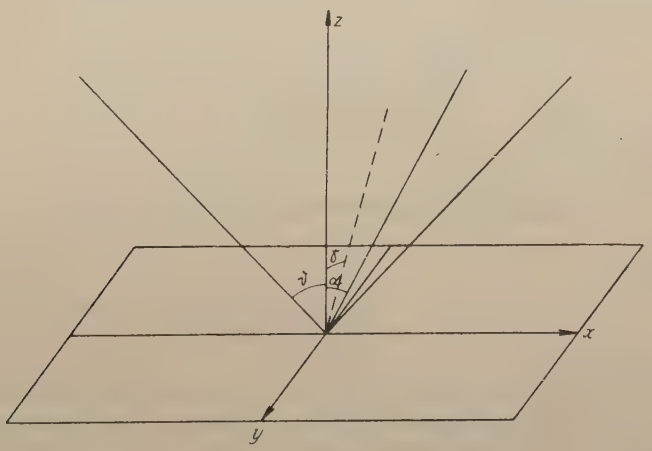

Fig. 6.

Assume the rotation axis of the crystal to coincide with the y-axis of the system of reference (Fig, 6), the reflecting plane of the crystal with the xy-plane, and the incident beam to lie in the xz-plane. Then a beam $\mathbf{s}$ reflected from a block of surface perpendicular to the axis also lies in the xz-plane. A beam $\mathbf{s}_1$ reflected from a block rotated by α about the y-axis lies in the xz-plane and yields, on the mobile film, a reflex removed by the angle α from that yielded by $\mathbf{s}$. For a block rotated by γ about the x-axis, the components of the unit vector of the incident beam $\mathbf{s}_{02}$ and reflected beam $\mathbf{s}_0$ are:

$$\mathbf{s}_{02}\,(-\cos(\vartheta+\varphi),\, 0,\, \sin(\vartheta+\varphi))$$

with ϑ denoting the angle fulfilling Bragg's equation,

$$\sin\varphi = \frac{\sin\vartheta\cos\vartheta}{\cos\gamma} - \frac{\sin\vartheta\cos\vartheta}{\cos\gamma}\left(1 - \frac{\sin^2\gamma}{\cos^2\vartheta}\right)^{1/2}$$

for small values of the angle γ

$$\sin\varphi = \tfrac{1}{2}\,\mathrm{tg}\,\vartheta\gamma^2$$

$$\mathbf{s}_2\left(\frac{(\cos^2\vartheta - \sin^2\gamma)^{1/2}}{\cos\gamma},\quad 2\sin\vartheta\sin\gamma,\quad \sin\vartheta\frac{\cos 2\gamma}{\cos\gamma}\right)$$

The beam $\mathbf{s}_2$ yields a reflex that is displaced by the angle $\Delta\vartheta$ with respect to that of $\mathbf{s}$ in the direction perpendicular to the line on the film (in the xz-plane):

$$\sin(\vartheta + \Delta\vartheta) = \frac{\sin\vartheta\cos 2\gamma}{\cos\gamma}(1 - 4\sin^2\vartheta\sin^2\gamma)^{-1/2}$$

hence, for small values of γ

$$\Delta\vartheta \approx (2\sin^2\vartheta - \tfrac{3}{2})\,\mathrm{tg}\,\vartheta\gamma^2 \tag{25}$$

The displacement along the line on the film (in the y-axis) is

$$\gamma' = \sin\vartheta\sin 2\gamma\,(\cos^2\vartheta - \sin^2\vartheta + \sin^2\vartheta\sin^2 2\gamma)^{-1}$$

at small values of γ

$$\gamma' = 2\,\mathrm{tg}\,\vartheta\gamma \tag{26}$$

By eq. (21), with a point focus X-ray tube, points of the line distant along the latter by $2\,\mathrm{tg}\,\vartheta\gamma$ are separated in the direction perpendicular to the line by the amount $2\,\mathrm{tg}^3\vartheta\gamma^2$. Finally, the widening of the line due to the presence of blocks rotated by the angle γ about the x-axis amounts to

$$\delta = (2\,\mathrm{tg}^3\vartheta + 2\,\mathrm{tg}\,\vartheta\sin^2\vartheta - \tfrac{3}{2}\,\mathrm{tg}\,\vartheta)\,\gamma^2 \tag{27}$$

If γ is of the same order of magnitude as the angle of maximum dispersion α_M of the blocks rotated about the y-axis, then δ is considerably smaller than α_M. Hence, the deformation of the line due to blocks rotated about the x-axis can be neglected.

A crystal block rotated by an angle φ about an axis lying in the reflecting xy-plane of the crystal and subtending the angle φ the with the y-axis (Fig 7) reflects equivalently to one rotated by α about the y-axis and by γ about the x-axis. The angles

ψ, φ, α and γ are related by the following formulas:

$$\sin \gamma = \sin \psi \sin \varphi \tag{28}$$

$$\sin \alpha = \sin \psi \cos \varphi \, (1 - \sin^2 \psi \sin^2 \varphi)^{-1/2} \tag{29}$$

for small values of the angle ψ

$$\gamma = \psi \sin \varphi \tag{30}$$

$$\alpha = \psi \cos \varphi \tag{31}$$

A block rotated by an angle ψ about an axis subtending an angle χ with the reflecting plane yields a reflex upon the film such as would result from a block rotated

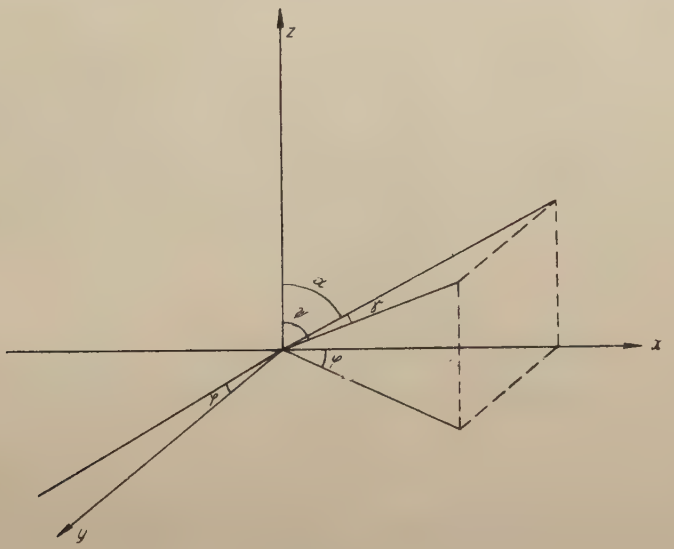

Fig. 7.

by the angle ψ_1 about the projection of the axis on the reflecting plane. The angles ψ, χ and ψ_1 (Fig. 8) are related by

$$\sin \frac{\psi_1}{2} = \cos \chi \sin \frac{\psi}{2} \tag{32}$$

at small values of ψ

$$\psi_1 = \psi \cos \chi$$

At small ψ a block rotated about an axis subtending the angle χ with the reflecting plane yields, on the mobile film, a reflex as would result from one rotated by α about the y-axis:

$$\alpha = \psi_1 \cos \varphi = \psi \cos \varphi \cos \chi = \psi \cos \Phi \tag{33}$$

φ denotes the angle subtended by the projection of the rotation axis of block on the reflecting plane, and the rotation axis of the crystal;
Φ denotes the angle between the rotation axis of the crystal block and that of the crystal (the y-axis).

The distribution function of the blocks, as computed from eqs. (14), (15), (16), (17), (18) or (22), (23), (24), results from the superposition of dispersion functions of blocks due to the different, variously oriented systems of dislocation lines. The distribution function of blocks due to a system of parallel edge dislocation lines is shrinked in the ratio cos Φ_e The angle Φ_e is the one subtended by the direction

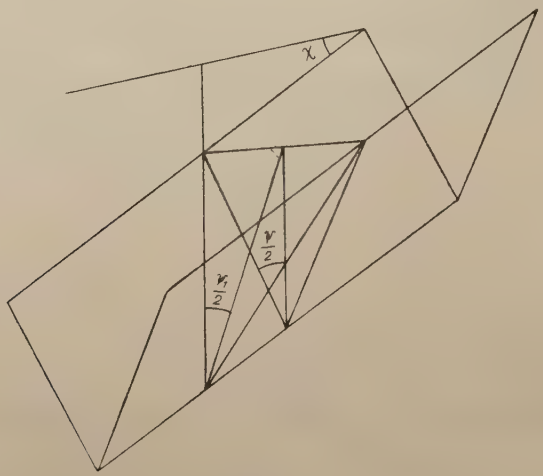

Fig. 8.

of the dislocation lines and the axis of rotation of the crystal. The distribution function of those due to screw dislocations is shrinked in the ratio cos Φ_n. The angle Φ_n is the one subtended by the normal to the plane of sliding containing the screw dislocation lines, and the rotation axis of the crystal.

9. Computation of Dislocation Density for Various Forms of the Block Distribution Function

If the block distribution function resulting from computation is of the Gauss curve form, then the mean angle between two neighbouring blocks can be computed from the formula (Gay, Hirsh Kelly 1953)

$$\alpha = \frac{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-p^2x^2} e^{-p^2y^2} |x-y|\, dx\, dy}{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-p^2(x^2+y^2)}\, dx\, dy} = \left(\frac{2}{\pi}\right)^{1/2} \frac{1}{p} = 0.479\,\beta \qquad (34)$$

wherein β is the half-width of the Gauss curve. Hence, by eq.(1), the density of dislocations at the block boundary amounts to

$$\varrho_{\text{bound}} = \frac{0.479\,\beta}{bt} \tag{35}$$

By eq. (2), the total dislocation density amounts to

$$\varrho = \left(\frac{0.479\,\beta}{b}\right)^2 \tag{36}$$

For a uniformly bent crystal, the block distribution curve $f(x)$ has the form a rectangle. The mean angle between neighbouring blocks is then

$$\alpha = \frac{2at}{L} \tag{37}$$

with $2a$ denoting the width of the function $f(x)$,
t — the block length, and
L — the length of the crystal surface scanned by the X-ray beam.
Assuming one dislocation to occur per inter-block boundary, the mean angle between the blocks is

$$\alpha = \frac{2ah}{L} \tag{38}$$

wherein h is the distance between the dislocations. Hence, the dislocation density in a uniformly bent crystal (Gay. Hirsh, Kelly 1953) amounts to

$$\varrho = \frac{2a}{Lb} \tag{39}$$

In practice, intermediate cases may occure, when, in addition to bending, there is dispersion of blocks in the crystal. It seems reasonable that, now, the resulting block distribution function $f(x)$ should be described by the folding of two functions, namely, the Gauss function $f_1(x)$ and a rectangular shape function $f_2(x)$. If

$$f_1(x) = e^{-p^2x^2}$$

$$f_2(x) = c \quad \text{for} \quad |x| < a, \quad f_2(x) = 0 \quad \text{for} \quad |x| > a, \quad c > 0$$

then $f(x)$ assumes the form

$$f(x) = \int_{-a}^{+a} c\, e^{-p^2(x-y)^2}\, dy = c \int_0^{a-x} e^{-p^2z^2}\, dz + c \int_0^{a+x} e^{-p^2z^2}\, dz \tag{40}$$

Hence, p and a can be computed.
The mean angle between contiguous blocks is now

$$\alpha = \left(\frac{2}{\pi}\right)^{1/2} \frac{1}{p} + \frac{2at}{L} \tag{41}$$

for $t = h$

$$\alpha = \left(\frac{2}{\pi}\right)^{1/2} \frac{1}{p} + \frac{2ah}{L} \tag{42}$$

Hence, the density at the block boundary is

$$\varrho_{\text{bound}} = \left(\frac{2}{\pi}\right)^{1/2} \frac{1}{pbt} + \frac{2a}{Lb} \tag{43}$$

The total density of dislocations amounts to

$$\varrho = \frac{1}{\pi p^2 b^2} + \frac{2a}{Lb} + \frac{2}{pb}\left(\frac{1}{4\pi^2 p^2 b^2} + \frac{a}{\pi Lb}\right)^{1/2} \tag{44}$$

Should the crystal contain no more than one system of edge dislocation lines parallel to its rotation axis, with Burgers' vector lying in the plane of reflexion, the density as computed from eqs. (35), (36), (39,) (43) and (44) would be numerically equal to the number of dislocation lines intersecting a unit area of the surface, the latter being perpendicular to the rotation axis of the crystal.

Should the crystal contain but one system of screw dislocstion lines lying in the plane of sliding, the latter being perpendicular to the rotation axis of the crystal, with Burgers' vector perpendicular to the reflecting surface, the density as computed form eqs. (35), (36), (39), (43) and (44) would be numerically equal to the number of dislocation lines intersecting the reflecting surface.

Nye (1953) proved that, for a crystal uniforly bent by a system of edge dislocation lines parallel to the bent surface, the curvature of the latter is given by the formula

$$\sigma = nb \cos \varepsilon \tag{45}$$

n denoting the number of dislocation lines intersecting a unit surface perpendiculary,
b — the module of Burgers' vector, and
ε — the angle between the bent surface and Burgers' vector.

In reality, a crystal exhibits a great variety of differently orientated systems of dislocation lines. By eqs. (33) and (35), the density as computed from eq. (39) can be written in the form of sums:

$$\varrho = \sum_{\nu=1}^{n_1} N_{e\nu} \cos \varepsilon_\nu \cos \Phi_{e\nu} + \sum_{\nu=1}^{n_2} \sum_{l=1}^{n_3} N_{s\nu} \cos \varepsilon_{n\nu} \cos \Phi_{nl\nu}$$

$N_{e\nu}$ is the number of edge dislocation lines subtending the angle $\Phi_{e\nu}$ with the rotation axis of the crystal and whose Burgers vector subtends the angle ε_ν with the plane of reflexion.

$N_{s\nu}$ is the number of screw dislocation lines in the plane of sliding whose normal subtends the angle $\Phi_{nl\nu}$ with the rotation axis of the crystal, and presenting a Burgers vector that subtends the angle $\varepsilon_{n\nu}$ with the normal to the plane of reflexion.

The density of eq (35) can be representend in the form of sums, as follows:

$$\varrho_{\text{bound}} = \sum_{\nu=1}^{n_1} \tfrac{1}{2} N_{e\nu} \cos \Phi_{e\nu} + \sum_{\nu=1}^{n_2} \sum_{l=1}^{n_3} N_{s\nu} \cos \varepsilon_{n\nu} \cos \Phi_{nl\nu}$$

The density as computed from eq. (36) asusmes the form

$$\varrho = [\sum_{\nu=1}^{n_1} (\tfrac{1}{2} N_{e\nu})^{1/2} \cos \Phi_{e\nu} + \sum_{\nu=1}^{n_2} \sum_{l=1}^{n_3} (N_{s\nu})^{1/2} \cos \varepsilon_{n\nu} \cos \Phi_{nl\nu}]^2$$

The density of eq. (43) assumes the form

$$\varrho_{\text{bound}} = \sum_{\nu=1}^{n_1} N_{e\nu} (\tfrac{1}{2} (1 - q_\nu) + q_\nu \cos \varepsilon_\nu)) \cos \Phi_{e\nu} + \sum_{\nu=1}^{n_2} \sum_{l=1}^{n_3} N_{s\nu} \cos \varepsilon_{n\nu} \cos \Phi_{nl\nu}$$

wherein q_ν is a quantity depending on the excess of dislocations of one sing for edge dislocations presenting a Burgers' vector that subtends the angle ε_ν with the plane of reflexion.

10. Conclusions

The method proposed yields the distribution function of mosaic blocks in a crystal, thus yielding the dislocation density therein.

1) In order to determine the block distribution function, it is necessary in experimental work that the dimensions of the X-ray tube focus, spectrograph slit and photometer slit should be chosen sufficiently small for the following conditions to be fulfilled:

a) $\alpha\beta \ll \alpha$

wherein α denotes the angle of dispersion of the mosaic blocks, and β — the angle of divergence of the beam incident on the crystal in the horizontal plane:

b) $\operatorname{tg} \varphi \cos^{-2} \varphi \Delta\varphi \ll \alpha$

with φ denoting the angle subtended by the beam incident on the film at the point under consideration on the line, and the one perpendicular to the axis of rotation of the crystal (Fig. 5). $\Delta\varphi$ is the angle of divergence in the vertical plane within the height of the photometer slit.

2) If experimental conditions were such that the immobile and the oscillating film were situated at one and the same distance from the crystal, then the block distribution function is yielded by computation from eqs. (14), (15), (16), (17) and (18). The formulas were derived on the assumption that the outer surface of the crystal alone reflects (absorption is very strong), and that inequalities of the surface within the limits set by the width of the beam can be neglected.

3) If experimental conditions were such that the immobile film was situated at one half of the distance from the oscillating film to the crystal, the block distribution function is obtained by computation from eqs. (22), (23) and (24). These

formulas also hold if there is reflexion from interior crystal faces (in crystals of low absorption); they eliminate error due to line deformation form inequalities of the surface within the beam width.

4) According to the form of the block distribution function obtained, the dislocation density at the boundary between the blocks is computed from eq. (36), (39) or (44).

The experimental curves required for computing the distribution function of mosaic blocks were obtained by measurement on several germanium monocrystals by Dr J. Auleytner and the present author, the numerical computations being carried out at the Institute of Computing Devices of the Polish Academy of Sciences (Zakład Aparatów matematycznych PAN).

The author wishes to thank Professor Dr L. Sosnowski for his valuable hints and discussions, and Dr J. Auleytner for suggesting the subject, and for his helpful advice

REFERENCES

Auleytner, J., *Acta phys. Polon.*, **17**, 111 (1958).
Auleytner, J., *Acta phys. Polon.*, **18**, 81 (1959*a*).
Auleytner, J., *Postępy Fizyki*, **10**, 77 (1959*b*).
Burgers, J. M., *Proc. Phys. Soc.* [London]. **52**, 23 (1940).
Gay, P., Hirsh, P. B., Kelly, A., *Acta Met.*, **1**, 315 (1953).
Nye, J. F., *Acta Met.*, **1**, 153 (1953).

ON THE DEPENDENCE OF THE PHOTON-ELECTRON RATIO ON THE DISTANCE FROM THE AXIS OF EXTENSIVE AIR SHOWERS OF COSMIC RADIATION

By A. Oleś

General Physics Laboratory, Academy of Mining and Metallurgy, Cracow

(Received November 11, 1959)

The transition curves were determined for two distances of the measuring tray from the detector of extensive air showers. The ratio of the number of photons to the number of electrons was found to increase as the mean distance from the axis of the showers recorded was augmented (from 20 to 100 m). For the range of distances applied, a ~ 50 % variation of the ratio resulted. The sign of the effect observed is in agreement with that predicted by theory of electron-photon cascade.

1. Introduction

The problem of the variation of the ratio of the number of photons to that of electrons (f/e) in an electron-photon cascade as a function of the distance from its axis was elaborated theoretically by Eyges and Fernbach [3] and Borsellino]2]. Their computations point to an increase in the excess of the photons over the electrons as the distance from the cascade axis increases. Notwithstanding the fact that the mean angle of deviation for particles of a given energy is larger for electrons than for photons, the circumstance that the latter are able to travel a long way through the atmosphere without loss of energy leads (for cascades of the age parameter $s > 0.6$) to an increase in the ratio f/e with increase of the distances from the axis of the cascade. Only young and narrow cascades with a small number of particles can exhibit the inverse effect. For these reasons, the foregoing conclusion of the electron-photon cascade theory should hold also with respect to extensive air showers of cosmic radiation.

Hitherto, measurements of the f/e ratio were carried out with the measuring tray located in the immediate neighbourhood of the detector of extensive showers. Milone [6] was the only author to investigate the dependence of the f/e ratio on the distance from the axis of the showers recorded. Milone carried out measurements for two positions of the measuring tray, the one at the detector, and another one

50 m therefrom. The result obtained by him was in either case $f/e \approx 1$. With respect to the small difference of the mean distances of either position from the axis of the showers, Milone was unable to detect the effect under consideration.

2. *Arrangement*

Measurements were carried out at Cracow on the terraces of the principal bulding of the Academy of Mining and Metallurgy. Fig. 1 shows the set of the trays of the Geiger-Müller counters. The detector (T) of extensive air showers consisted of three trays A, B, C of an area of $S = 0.45$ m² each. A fourth, measuring tray (G) remained

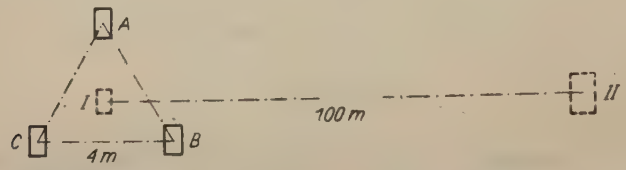

Fig. 1. Diagram of device. *ABC* — detector of extensive air showers. I, II — two positions of measuring tray.

in coincidence with the detector, and was situated directly beneath an absorbent of variable thickness of 0 to 10 mm Pb. Measurements were carried out for two positions of the measuring tray: position I — within the triangle determined by the detector, and II — at a distance of 100 m therefrom. In the former case the area of tray G was $\Sigma = 0.28$ m², and in the latter $\Sigma = 0.91$ m². Aluminium counters, of a wall thickness of 1 mm and a threshold energy of ~ 1 MeV for electron recorded, were used.

3. *Experimental Results*

The numbers of coincidences registered are assembled in Tables I and II.

Table I

Results of measurements with tray G at position I.

mm Pb	T	G	G/T	t
0	16405	9192	0.560	$77^h 46^m$
3	8522	4889	0.574	38 12
5	10555	6176	0.585	52 47
7	9220	5347	0.580	44 20
10	10715	5984	0.559	50 36

Table II

Results of measurements with tray G at position II.

mm Pb	T	G	G/T	t
0	16579	2269	0.137	$78^h 33^m$
2	14189	2203	0.155	71 27
5	23998	4290	0.179	113 16
7	11294	1891	0.167	56 26
10	23400	3666	0.157	105 13

Fourfold coincidences $ABCG$ can by expessed by formula

$$G(t) = k \int_0^\infty (1 - e^{-xS})^3 \, (1 - e^{-r R(t) x S}) \, x^{-(\gamma+1)} \, dx \qquad (1)$$

wherein $kx^{-(\gamma+1)}$ is the differential density spectrum of the showers recorded, S — the area of the detector tray, $r = z \, \Sigma/S$ with z denoting the ratio of the mean densities of the particles over the measuring tray to that over the detector, and Σ — the area of the measuring tray G; $R(t)$ is given by

$$R(t) = P_e(t) + f/e \, P_f(t) \qquad (2)$$

wherein $P_e(t)$ and $P_f(t)$ are the probabilities that, an electron or photon respectively falling on the absorber of a given thickness t will produce beneath it at least one secondary electron which operates the counter tray covered with that absorber; f/e is the photon/electron ratio required, determined for energies of particles of either kind exceeding the threshold energy of the electron recording device.

By eq. (1), $R(t)$ can be determined from the experimental data. The transition curves obtained are shown in Fig. 2. These served subsequently for finding the f/e ratio. In the computations (wherein eq. (2) was used), the experimental value of $P_f(t)$ was taken from the paper by Massalski and Oleś [5], and $P_e(t)$ was assumed as computed by Arley [1]. The following values resulted for the f/e ratio:

with the measuring tray at position I — $(f/e)_I = 0.95 \pm 0.10$
with the measuring tray at position II — $(f/e)_{II} = 1.40 \pm 0.15$

In an earlier paper [5], photons (especially of energy lower than the threshold energy of the electron recording device) were found to affect the coincidence frequency of the tray when not shielded by the absorbent. The effect involves an increase in the $ABCD$ coincidence frequency for the uncovered tray, leading to a lowering of the maximum of the transition curve and yielding a lower value of the f/e ratio. If the respective correction is accounted for, the values for $(f/e)_I$ and $(f/e)_{II}$ are raised by about 20 %.

The mean distance of the measuring tray at position I from the axis of the showers recorded amounts to about 20 m. The analogous distance was evaluated also for position II assuming the lateral distribution function for the particle density according to Nishimura and Kamata [4]. Computation yields for the measuring tray at position II a mean distance of $\sim$ 100 m from the axis of the showers recorded.

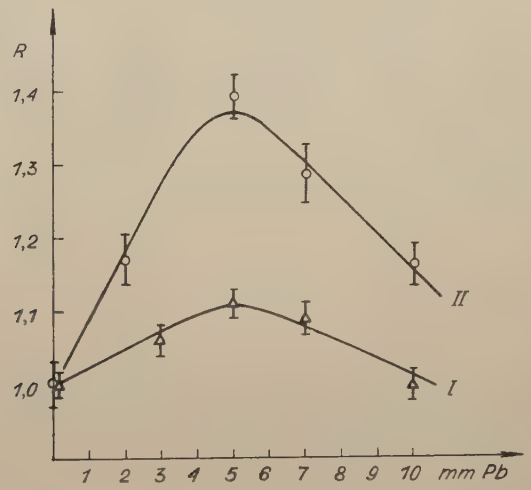

Fig. 2. Experimental transition curves. Curve I was obtained with tray situated within extensive air shower detector. Curve II — with tray 100 m distant from detector.

4. *Conclusions*

The experimental variation of the ratio of the numbers of photons to the electrons in extensive air showers of cosmic radiation as a function of the distance from the axis of the showers recorded exhibits a trend in agreement with predictions of the electron-photon cascade theory.

According to experimental data presented in Greisen's article [4] the lateral distribution of electron density in extensive air showers at sea level is in agreement with the distribution for an electron-photon cascade of the age parameter $s = 1.2$. Similarly, the lateral distribution of photon density of extensive air showers be expected to resemble the one in a single electron-photon cascade of the above value of the parameter s. Hence in extensive showers, the value of f/e should be expected to increase as the distance from the axis augments (as is the case in an electron-photon cascade of $s = 1.2$).

With the device described above, a ratio of $f/e \approx 1$ is obtained for the position of the measuring tray within extensive air shower detector. This, together with the increase of the f/e ratio as the distance of the measuring tray from the detector augments, leads to the conclusion that the photon/electron ratio throughout the entire shower exceeds 1.

The author wish to thank Professor Dr M. Mięsowicz and Docent Dr J. Massalski for their valuable advice and discussions; Eng. T. Owsiak and the technical team of our Institute for their help in the construction of the measuring device.

REFERENCES

[1] Arley, N., *On the theory of the stochastic processes*, Copenhagen, 1948.
[2] Borsellino, A., *Nuovo Cimento*, **7**, 323 and 601 (1950).
[3] Eyges, L., Fernbach, S., *Phys. Rev.*, **82**, 23 (1951).
[4] Greisen, K., *Progress in Cosmic Ray Physics*, edited by Wilson, J. G., Amsterdam, 1956.
[5] Massalski, J., Oleś, A., *Acta phys. Polon.*, **17**, 401, (1958).
[6] Milone, C., *Nuovo Cimento*, **9**, 549 (1952) and **10**, 1126 (1953).

ON THE STATE OF A FERMI-SYSTEM WITH CORRELATION OF PAIRS OF PARTICLES WITH PARALLEL SPINS. I. GROUND — STATE.

By Zygmunt Galasiewicz[1]

Joint Institute for Nuclear Research — Dubna, USSR

(Received December 9, 1959)

The solution of a compensation equation of dangerous graphs, for a system of interacting Fermi-particles, as an odd function of wave vectors is investigated. It is shown that this solution leads to the anomalous state the energy of which is smaller than the energy of the normal state. The energy difference is equal to the binding energy of the pairs of fermions with parallel spin-moments. This anomalous state can exist in the presence of attractive Coulomb interactions. The elementary excitations are separated from the ground state by a gap the width of which depends on the direction of the wave vector. For the direction perpendicular to the spin direction, energy gap is zero and therefore the state cannot be superconducting. The binding energy of the pairs and the width of the energy gap are of the same order as in the theory based on the creation of pairs with antiparallel spin-moments, where in the expansion of the interaction term in spherical harmonics, the first coefficient is more than three times greater than the zero coefficient. In the absence of the magnetic field the total spin of a system is zero. In the presence of the magnetic field the total spin differs from zero. The paramagnetic susceptibility is calculated.

1. Introduction

Consider a dynamical system of Fermi-particles with the Hamiltonian

$$H = \sum_{f,f'} T(f,f')\, a_f^+ a_{f'} + \tfrac{1}{2} \sum_{f_1,f_2,f_1',f_2'} U(f_1,f_2;f_2',f_1')\, a_{f_1}^+ a_{f_2}^+ a_{f_2'} a_{f_1'}, \tag{1}$$

$$T(f,f') = I(f,f') - \lambda\delta(f-f'),$$

where I is the Hamiltonian of a particle, U — the interaction energy, λ — the chemical potential, a_f, a_f^+ the Fermi-amplitudes and f is a set of indices characterizing one-particle states. In our case $f = (\vec{p}, \sigma_z)$, where $\vec{p}$ is a wave vector and σ_z a spin index.

---

[1]) Permanent adress: Institute of Theoretical Physics, Univesity of Wrocław, Wrocław, Poland

Similarly as in [1] (and also in [2] and [3]), we transform the Hamiltonian (1) by means of the transformation

$$a_f = \sum_\nu (u_{f\nu}\alpha_\nu + v_{f\nu}\alpha_\nu^+). \tag{2}$$

To secure the canonical character of the transformation (2) the functions $\{u, v\}$ must be connected by the orthonormality relations

$$\sum_\nu \{u_{f\nu} u^*_{f'\nu} + v_{f\nu} v^*_{f'\nu}\} = \delta(f - f')$$

$$\sum_\nu \{u_{f\nu} v_{f'\nu} + u_{f'\nu} v_{f\nu}\} = 0 \tag{3}$$

We find the functions $\{u, v\}$ from the additional equations obtained from the compensation principle of dangerous graphs (see [2])

$$<\alpha_{\nu_1}\alpha_{\nu_2} H>_0 = 0 \tag{4}$$

(the expectation value corresponds to the vaccum state in the α-represention) which are equivalent to the following equations for the functions $F(f, f')$, $\Phi(f, f')$:

$$\sum_f \{\vartheta(f_1, f)\,\Phi(f, f_2) + \vartheta(f_2, f)\,\Phi(f_1, f)\} + S(f_1, f_2) -$$

$$- \sum_f \{F(f, f_1)\,S(f, f_2) + F(f, f_2)\,S(f_1, f)\} = 0, \tag{5}$$

where

$$F(f, f') = \sum_\nu v^*_{f\nu} v_{f'\nu}, \qquad \Phi(f, f') = -\Phi(f', f) = \sum_\nu u_{f\nu} v_{f'\nu} \tag{6}$$

and

$$S(f_1, f_2) = \sum_{f_1', f_2'} U(f_1, f_2; f_2', f_1')\,\Phi(f_1', f_2'),$$

$$\vartheta(f_1, f) = T(f_1, f) + \sum_{f', f''} \{U(f_1, f''; f', f) - U(f_1, f''; f, f')\}\,F(f'', f') \tag{7}$$

In order to represent the functions F, Φ in the form (6) these functions must satisfy the additional subsidiary conditions (see [2])

$$F(f_1, f_2) = \sum_f \{F(f_1, f)\,F(f, f_2) + \Phi(f, f_1)\,\Phi(f, f_2)\} \tag{8a}$$

$$\sum_f \{F(f_1, f)\,\Phi^*(f, f_2) + F(f_2, f)\,\Phi^*(f, f_1)\} = 0 \tag{8b}$$

From formula (1) of paper [2] we know that

$$T(f, f') = (E(p) - \lambda)\,\delta(f - f') = \varepsilon(p)\,\delta(f - f'),$$

$$U(f_1, f_2; f_2' f_1') = \frac{1}{V} J(\vec{p}_1, \vec{p}_2; \vec{p}_2', \vec{p}_1')\,\delta(\vec{p}_1 + \vec{p}_2 - \vec{p}_1' - \vec{p}_2')\,\delta(\sigma_1 - \sigma_1')\,\delta(\sigma_2 - \sigma_2'),$$

$$J(\vec{p}_1, \vec{p}_2; \vec{p}_2', \vec{p}_1') = J(\vec{p}_2, \vec{p}_1; \vec{p}_1', \vec{p}_2') = J(-\vec{p}_1, -\vec{p}_2; -\vec{p}_2' - \vec{p}_1'). \tag{9}$$

The solution of equation (5) with the subsidiary conditions (8a, b) may be put in the form

$$F(f,f') = F(\vec{p})\,\delta(f-f'), \Phi(f,f') = \Phi(f)\,\delta(f+f'),$$
$$\Phi(\vec{p},+) = -\Phi(\vec{p}), \qquad \Phi(\vec{p},-) = \Phi(\vec{p}). \tag{10}$$

The functions $F(\vec{p})$ and $\Phi(\vec{p})$ are assumed to be real, invariant under the transformation of momentum reflection, and to satisfy the following equations

$$2\mathfrak{z}(p)\,\Phi(\vec{p}) + \frac{1-2F(\vec{p})}{V}\sum_{\vec{p}'} J(\vec{p},-\vec{p};-\vec{p}',\vec{p}')\,\Phi(\vec{p}') = 0$$

$$F(\vec{p}) = F^2(\vec{p}) + \Phi^2(\vec{p}), \tag{11}$$

where

$$\mathfrak{z}(p) = \varepsilon(p) - \frac{1}{V}\sum_{\vec{p}}\left[2J(\vec{p},\vec{p}';\vec{p}',\vec{p}) - J(\vec{p},\vec{p}';\vec{p},\vec{p}')\right]F(\vec{p}'). \tag{12}$$

The solution of this equations leads to the usual formulae of the theory of superconductivity.

2. *The solution of the compensiation relation in the case $\Phi(\vec{p},\sigma_z)$ is an odd function of $\vec{p}$*

From the form equations (11) we see that a solution of this equations exist also in the case when $\Phi(\vec{p},\sigma_z)$ is an odd function of $\vec{p}$. Instead of (10), we put therefore

$$F(f,f') = F(\vec{p})\,\delta(f-f'),\ \Phi(f,f') = \Phi(f)\,\delta(\vec{p}+\vec{p}')\,\delta(\sigma-\sigma'),$$
$$F(\vec{p}) = F(-\vec{p}).\ \Phi(\vec{p},\sigma_z) = -\Phi(\vec{p},-\sigma_z),\ \Phi(\vec{p},\sigma_z) = -\Phi(-\vec{p},\sigma_z). \tag{13}$$

The functions of the form

$$\Phi(\vec{p},\sigma_z) = e_z\,\sigma_z\sum_{sz} S_{\sigma_z}(s_z)\,\varphi(\vec{p}) = \sigma_z\cos\vartheta\,\varphi(\vec{p}),\ \varphi(\vec{p}) = \varphi(-\vec{p}) \tag{13'}$$

satisfy these conditions, where $\vec{e} = \vec{p}/p$. $S_{\sigma_z}(s_z)$ is a spin function ($S_{\sigma_z}(\sigma_z) = 1$, $S_{\sigma_z}(s_z) = 0$ for $\sigma_z \neq s_z$, $\hat{\sigma}_z S = \sigma_z S$, $\sigma_z = \pm 1$), and ϑ is the angle between the wave vector and the quantization axis of the electron spins. Substituting (13) in (5) and (8a, b), we get for the functions $F(\vec{p})$, $\Phi(\vec{p},\sigma_z)$ the equations (11) (from now on we shall denote the spin index simply by σ). We define

$$C(\vec{p},\sigma) = -\frac{1}{V}\sum_{\vec{p}'} J(\vec{p},-\vec{p};-\vec{p}',\vec{p}')\,\Phi(\vec{p}',\sigma) \tag{14}$$

$C(\vec{p},\sigma)$ is an odd function of $\vec{p}$ and σ

$$C(\vec{p},\sigma) = -C(-\vec{p},\sigma),\ C(\vec{p},\sigma) = -C(\vec{p},-\sigma),\ C^2(p,\sigma) = C^2(\vec{p}). \tag{15}$$

From equations (11) we obtain

$$\Phi(\vec{p}, \sigma) = \frac{C(\vec{p}, \sigma)}{2\sqrt{C^2(\vec{p}) + \vartheta^2(p)}}, \qquad F(\vec{p}) = \frac{1}{2}\left\{1 - \frac{\vartheta(p)}{\sqrt{C^2(\vec{p}) + \vartheta^2(p)}}\right\}.$$

hence

$$C(\vec{p}, \sigma) = -\frac{1}{2V}\sum_{\vec{p}'} J(\vec{p} - \vec{p}; -\vec{p}', \vec{p}')\,\frac{C(\vec{p}', \sigma)}{\sqrt{C^2(\vec{p}) + \vartheta^2(p)}}. \tag{17}$$

We take

$$J(\vec{p}, -\vec{p}; -\vec{p}', \vec{p}') = J(|\vec{p} - \vec{p}'|) = \sum_n J_n(\vec{p}, \vec{p}')\, P_n(\cos\gamma) \tag{18}$$

where γ is the angle between the vectors $\vec{p}$ and $\vec{p}'$

$$\cos\gamma = \cos\vartheta\cos\vartheta' + \cos\varphi\sin\vartheta\,\mathrm{son}\,\vartheta' \tag{19}$$

ϑ, ϑ' are the angles between the vectors $\vec{p}$, $\vec{p}'$ and the polar axis, and φ is the angle round the polar axis. According to the addition theorem, we have

$$P_n(\cos\gamma) = P_n(\cos\vartheta)\,P_n(\cos\vartheta') + 2\sum_{m/1}^{n}\frac{(n-m)!}{(n+m)!}\,P_n^m(\cos\vartheta)\,P_n^m(\cos\vartheta')\cos m\varphi. \tag{20}$$

Substituting (20) into (17) and writing the result in integral form, we obtain

$$C(\vec{p}, \sigma) = -$$
$$-\frac{1}{2}\frac{1}{(2\pi)^2}\sum_n P_n(\cos\vartheta)\int_0^\infty\!\!\int_0^\pi J_n(p, p')\,P_n(\cos\vartheta')\,\frac{C(p', \cos\vartheta', \sigma)\,p^2\,dp'}{\sqrt{C^2(p', \cos\vartheta') + \vartheta^2(p')}}\sin\vartheta'\,d\vartheta'. \tag{21}$$

The solution of (21) may be put in the form

$$C(\vec{p}, \sigma) = \sigma\cos\vartheta\Psi(p). \tag{22}$$

Moreover, we consider the case when the interaction is effective only for the momentum p near p_F and take into account only $J_1(p, p') \approx J_1(p_F)$ (the contribution from $J_0(p, p')$ vanishes). For the radial part of $C(\vec{p}, \sigma)$ we get the equation

$$\Psi(p) = -\frac{1}{2}\frac{p_F^2 J_1}{(2\pi)^2}\int_{-1}^{1} x^2\,dx\int_{p_F-\Delta}^{p_F+\Delta}\frac{\Psi(p')\,dp'}{\sqrt{\Psi^2(p')\,x^2 + \zeta^2(p')}}. \tag{23}$$

Looking for a solution which changes slowly with p in the neighbourhood of p_F and taking $\vartheta(p) \approx E'(p - p_F)$, we get

$$\Psi = 2\,\omega\,e^{1/\varrho_1}\,e^{3/\varrho_1} = 2\,\omega\,e^{1/\varrho_1}\,e^{-3/\varrho_1}, \tag{24}$$

where

$$-\tilde{\varrho}_1 = -\frac{dn}{dE} J_1 = \varrho_1 > 0, \qquad \frac{dn}{dE} = \frac{p_F^2}{2\pi^2 E'}, \qquad \omega = E'\Delta$$

(we have to replace the interaction of electrons with sound quanta by an effective electron-electron interaction, which is different from zero in a narrow layer near the Fermi surface $E_F \pm \omega$). In order that $\Psi \to 0$ for $\tilde{\varrho}_1 \to 0$ the effective interaction must be attractive ($J_1(p_F) < 0$). From the solution of (23), we get the function $C(\vec{p}, \sigma)$ and consequently the functions $F(\vec{p})$ and $\Phi(\vec{p}, \sigma)$. Moreover, we can obtain another solution of (11) by putting $\Phi(p, \sigma) = 0$ and $F(\vec{p}) = 1$ for $p < p_F$ and $F(\vec{p}) = 0$ for $p > p_F$.

We examine now the special form assumed by the transformation (2) when the functions F and Φ are given by (13). Taking into account (6) and (13'), we see that

$$u_{fv} = u(\vec{p})\,\delta(f - v), \quad v_{fv} = v(-\vec{p}, \sigma)\,\delta(\vec{p} + \vec{l})\,\delta(\sigma - s), \quad f = (\vec{p}, \sigma), \quad v = (\vec{l}, s)$$

$$v(\vec{p}, \sigma) = -v(-\vec{p}, \sigma), \; v(\vec{p}, \sigma) = -v(\vec{p}, -\sigma), \; v(\vec{p}, \sigma) = \sigma \cos\vartheta\, w(\vec{p}), \; w(\vec{p}) = w(-\vec{p}) \tag{25}$$

and

$$F(\vec{p}) = v^2(\vec{p}, \sigma) = \cos^2\vartheta\, w^2(\vec{p}), \qquad \Phi(\vec{p}, \sigma) = u(\vec{p})\, v(\vec{p}, \sigma). \tag{26}$$

Therefore the transformation (2) has now the form

$$a_{\vec{p}\sigma} = u(\vec{p})\,\alpha_{\vec{p}\sigma} - v(\vec{p}, \sigma)\,\alpha^+_{-\vec{p}\sigma}, \quad a_{-\vec{p}\sigma} = u(\vec{p})\,\alpha_{-\vec{p}\sigma} + v(\vec{p}, \sigma)\,\alpha^+_{\vec{p}\sigma}. \tag{27}$$

If by means of (27) we transform the Hamiltonian

$$H = \sum_{\vec{p}, \sigma} \varepsilon(p)\, a^+_{\vec{p}\sigma}\, a_{\vec{p}\sigma} + \frac{1}{2V} \sum_{\vec{p}_1 + \vec{p}_2 = \vec{p}_1' + \vec{p}_2'} J(\vec{p}_1, \vec{p}_2; \vec{p}_2', \vec{p}_1')\, a^+_{\vec{p}_1\sigma_1}\, a^+_{\vec{p}_2\sigma_2}\, a_{\vec{p}_2'\sigma_2}\, a_{\vec{p}_1'\sigma_1} \tag{28}$$

(see (1) and (9)), we see that we must write the compensation equation for the dangerous diagrams for the pairs of fermions with parallel spins and antiparallel momenta.

For the functions $u(\vec{p})$, $v(\vec{p}, \sigma)$ we get a first solution for the normal state, namely

$$u(\vec{p}) = \Theta_G(\vec{p}) = \begin{cases} 1, & E(\vec{p}) > E_F \\ 0, & E(\vec{p}) < E_F \end{cases}, \qquad v(\vec{p}, \sigma) = \Theta_F(\vec{p}, \sigma) = \begin{cases} 0, & E(\vec{p}) > E_F \\ \pm 1, & E(\vec{p}) < E_F \end{cases} \tag{29}$$

(where $\Theta_F(\vec{p}, \sigma) = -\Theta_F(-\vec{p}, \sigma), \ldots$), and a second solution by taking into account (26) and (11), namely

$$u^2(\vec{p}) = \frac{1}{2}\left\{1 + \frac{\vartheta(p)}{\sqrt{\Psi^2 \cos^2\vartheta + \vartheta^2(p)}}\right\}, \qquad v^2(\vec{p}, \sigma) = \frac{1}{2}\left\{1 - \frac{\vartheta(p)}{\sqrt{\Psi^2 \cos^2\vartheta + \vartheta^2(p)}}\right\} \tag{30}$$

This solution is of the type obtained by Bogolubov [1] but it depends on the direction of the vector $\vec{p}$ and for $\vartheta = \pi/2$ it goes over into the solution for the normal state.

3. *The energy difference for two states and the energy gap*

Now we find the mean energy in the vacuum state in α-representation and evaluate the difference for solutions (29) and (30). The general formula for the mean value is

$$E = <H>_0 = \tfrac{1}{2} \sum_{f,f'} [T(f,f') + \mathfrak{z}(f,f') + \\ + \tfrac{1}{2} \sum_{f_1,f_2,f'_1,f'_2} U(f_1,f_2;f'_2,f'_1)\, \Phi^*(f_1,f_2)\, \Phi(f'_1,f'_2) \tag{31}$$

Taking into account (9) and (13), we obtain

$$<H>_0 = \sum_{\vec{p}} [\varepsilon(p) + \mathfrak{z}(p)]\, F(\vec{p}) + \frac{1}{V} \sum_{\vec{p}\,\vec{p}'} J(\vec{p}, (-\vec{p}; -\vec{p}', \vec{p}')\, \Phi(\vec{p}, \sigma)\, \Phi(\vec{p}', \sigma)$$

$$\cong 2 \sum_{\vec{p}} \mathfrak{z}(p)\, F(\vec{p}) - \sum_{\vec{p}} \frac{C^2(\vec{p})}{2\sqrt{C^2(\vec{p}) + \mathfrak{z}^2(p)}} \tag{32}$$

The energy difference for the two solutions may be found with the help of the identity

$$v^2(\vec{p}, \sigma) - \Theta_F^2(\vec{p}) =$$

$$= \frac{\Theta_G(\vec{p})}{2} \left\{ 1 - \frac{\mathfrak{z}(p)}{\sqrt{\Psi^2 \cos^2\vartheta + \mathfrak{z}^2(p)}} \right\} - \frac{\Theta_F^2(p)}{2} \left\{ 1 + \frac{\mathfrak{z}(p)}{\sqrt{\Psi^2 \cos^2\vartheta + \mathfrak{z}^2(p)}} \right\} \tag{33}$$

Finally we get

$$\frac{\Delta \bar{E}}{V} = \frac{2}{V} \sum_{\vec{p}} \mathfrak{z}(p) [v^2(\vec{p}, \sigma) - \vartheta_F^2(\vec{p})] - \frac{1}{2V} \sum_{\vec{p}} \frac{C^2(\vec{p})}{\sqrt{C^2(\vec{p}) + \mathfrak{z}^2(p)}} =$$

$$= -\frac{2}{3} \frac{dn}{dE}\, e^{1/3}\, \omega^2\, (e^{-2/\varrho_1})^3. \tag{34}$$

From this it is clear that the solution (30) leads to an energy state which lies lower than the normal state. This is also the anomalous state connected with the pairs of fermions with parallel spin-moments.

Let us compare the ground state for the case of pairs of fermions with antiparallel spins with the ground state for the case of pairs with parallel spins. We see that the binding energy for the pairs of electrons with parallel spin-moments is proportional to $[\exp(-2/\varrho_1)]^3$, while the binding energy for the pairs of electrons with antiparallel spin-moments is proportional to $\exp(-2/\varrho_0)$. When the case $J_1(p,p') > 3\, J_0(p,p')$

is not realized the ground state of the system for pairs of fermions with antiparallel spin-moments has α lower energy than the state for pairs with parallel spin-moments.

We shall now find the formula for the elementary excitation in the superconducting state. The general formula is given by [3]

$$<\alpha_\nu|H|\alpha_\nu^+>_0 = \sum_{f,f'} \mathfrak{z}(f,f')\,(u^*_{f\nu}u_{f'\nu} - v^*_{f\nu}v_{f'\nu}) + \\ + \tfrac{1}{2}\sum_{f_1,f_2,f_1',f_2'} U(f_1,f_2;f_2',f_1')\,[\Phi(f_2,f_1)\,(u^*_{f_1'\nu}v^*_{f_2'\nu} - v^*_{f_1'\nu}u_{f_2'\nu}) + \\ + \Phi^*(f_2,f_1)\,(u_{f_1'\nu}v_{f_2'\nu} - u_{f_2'\nu}v_{f_1'\nu})]. \tag{35}$$

By the way we remark that (35) is identical with the formula for the energy of elementary excitations given by Landau in [4]:

$$\Omega(\nu) = \frac{\delta\bar{E}}{\delta n_\nu} \tag{36}$$

Here E is the diagonal matrix element of the Hamiltonian (1) transformed by means of transformation (2) of [3] when the expectation value of $\alpha_\nu^+\alpha_\nu$ is not zero but n_ν. In this case the mean energy is given by a formula of the form (31) but the functions F and Φ are now of the form (see [2])

$$\Phi(f_1,f_2) = \sum_\nu \{u_{f_1\nu}v_{f_2\nu}(1-n_\nu) + v_{f_1\nu}u_{f_2\nu}n_\nu\},$$

$$F(f_1,f_2) = \sum_\nu \{v^*_{f_1\nu}v_{f_2\nu}(1-n_\nu) + u^*_{f_1\nu}u_{f_2\nu}n_\nu\}. \tag{37}$$

For $T=0$, that is $n_\nu = 0$, formulae (37) pass over into (6). Taking also into account that in (31) F and Φ are functions of n_ν; we get

$$\frac{\delta\bar{E}}{\delta n_\nu} = \frac{1}{2}\sum_{f,f'}\left\{\frac{\delta\mathfrak{z}(f,f')}{\delta n_\nu}F(f,f') + [T(f,f') + \mathfrak{z}(f,f')]\frac{\delta F(f,f')}{\delta n_\nu}\right\} + \\ + \frac{1}{2}\sum_{f_1,f_2,f_1',f_2'} U(f_1,f_2;f_2',f_1')\left[\frac{\delta\Phi^*(f_1,f_2)}{\delta n_\nu}\Phi(f_1',f_2') + \Phi^*(f_1,f_2)\frac{\delta\Phi(f_1'f_2')}{\delta n_\nu}\right] \tag{38}$$

Finding $\frac{\delta F}{\delta n_\nu}$, $\frac{\delta\Phi}{\delta n_\nu}$ from (37), we see that

$$\left(\frac{\delta\bar{E}}{\delta n_\nu}\right)_{n_\nu=0} = <\alpha_\nu|H|\alpha_\nu^+>_0 = \Omega(\nu). \tag{39}$$

Introducing (13) into (35), we obtain

$$\Omega(\vec{p}) = \mathfrak{z}(p)\,[u^2(\vec{p}) - v^2(\vec{p},\sigma)] - \frac{2}{V}u(\vec{p})\,v(\vec{p},\sigma)\sum_{\vec{p}'}\mathrm{J}(\vec{p},-\vec{p};-\vec{p}',\vec{p}')\,v(\vec{p}',\sigma)\,u(\vec{p}). \tag{40}$$

From (11) and (26), we get

$$\Omega(\vec{p}) = \sqrt{\Psi^2 \cos^2 \vartheta + \vartheta^2(p)} \tag{41}$$

On the Fermi-surface

$$\Omega(p, \cos\vartheta) = |\cos\vartheta|\, 2\omega\, e^{+\frac{1}{2}}\, e^{-2/\varrho_1} = \Delta \tag{42}$$

Thus the fermion excitations in the anomalous state are separated from the energy of the ground state by the gap Δ, depending on the direction of the vector $\vec{p}$. Since this energy gap can be equal to zero, no current-carrying state of the system will exist with respect to weak pertubations. This means that the anomalous state considered by us is not superconducting.

4. The influence of the electron-electron Coulomb repulsion

Let us now examine the combined effect of the electron-phonon interaction and the Coulomb interaction between the electrons (see [1] and [5]). Similarly as in the case of the electron-phonon interaction (which is essential only near the Fermi--surface, $E_F \pm \omega$), we replace the Coulomb interaction by a model interaction. Because of the screening of the Coulomb interaction, we replace it by a constant attraction of electrons (essential only near the Fermi-surface $E_F \pm \omega_1$, $\omega_1 > \omega$). Thus we put

$$J_1(\zeta, \zeta') = \begin{cases} J_{1\mathrm{ph}}(\vartheta, \vartheta') + J_{1C}(\vartheta, \vartheta'), & |\vartheta|, |\vartheta'| < \omega \\ J_{1C}(\vartheta, \vartheta') & |\vartheta| \text{ or } |\vartheta'| > \omega \end{cases} \tag{43}$$

We obtain now equations analogous to (23). from which we can determine $\Psi(\vartheta)$ in the two intervals: $|\vartheta| < \omega$ and $\omega < |\vartheta| < \omega_1$

$$\Psi(\vartheta) = -\frac{1}{4}\frac{dn}{dE}\int_{-1}^{1} x^2\, dx \int_{-\omega_1}^{\omega_1} J_1(\vartheta, \vartheta')\, \frac{\Psi(\vartheta')\, d\vartheta'}{\sqrt{\Psi^2(\vartheta')\, x^2 + \vartheta'^2}}\,. \tag{44}$$

We put

$$\Psi(\vartheta) = \begin{cases} \Psi_0, & |\vartheta| < \omega \\ \Psi_1, & \omega < |\vartheta| < \omega_1 \end{cases} \tag{45}$$

From (43) we get

$$\begin{aligned} -\Psi_0 &= \frac{1}{3}\left[\varrho_{1C}\Psi_1 \ln\frac{\omega_1}{\omega} + \Psi_0(\tilde{\varrho}_{1\mathrm{ph}} + \varrho_{1C}) \ln\frac{2\omega}{\Psi_0} + \frac{1}{3}\Psi_0(\tilde{\varrho}_{1\mathrm{ph}} + \varrho_{1C})\right], \\ -\Psi_1 &= \frac{1}{3}\left[\varrho_{1C}\Psi_1 \ln\frac{\omega_1}{\omega} + \Psi_0\varrho_1 \ln\frac{2\omega}{\Psi_0} + \frac{1}{3}\Psi_0\varrho_{1C}\right], \end{aligned} \tag{46}$$

where

$$-\tilde{\varrho}_{1\mathrm{ph}} = -J_{1\mathrm{ph}}\frac{dn}{dE} = \varrho_{1\mathrm{ph}} > 0, \quad \varrho_{1C} = J_{1C}\frac{dn}{dE} > 0$$

Hence

$$\frac{1}{3}\left(\ln\frac{2\omega}{\Psi_0}+\frac{1}{3}\right)\left(\varrho_{1\,\mathrm{ph}}-\frac{\varrho_{1C}}{1+\frac{1}{3}\varrho_{1C}\ln\frac{\omega_1}{\omega}}\right)=1 \tag{47}$$

Thus the condition of the appearing of a supereonducting state the for case of pairs of fermions with antiparallel spin-moments has an analogous form to that in [1]. It reads

$$\varrho_{1\,\mathrm{ph}}>\frac{\varrho_{1C}}{1+\frac{1}{3}\varrho_{1C}\ln\frac{\omega_1}{\omega}} \tag{48}$$

5. *Total spin of a system and paramagnetic susceptibility*

The spin moment for unit volume is given by the formula

$$\tilde{S}^z=\frac{1}{V}\sum_{\vec{p}}\{F(\vec{p},+)-F(\vec{p},-)\}. \tag{49}$$

We pass over to the operators α and find (49) for the state C_0 (vacuum state in the α representation)

$$S^z=\frac{1}{V}\sum_{\vec{p}}\{F(\vec{p},+)-F(\vec{p},-)\}. \tag{50}$$

In the absence of a magnetic field $F(\vec{p},+)=F(\vec{p},-)$ and therefore $S^z=0$.

Consider now a system of Fermi-particles in the presence of an external, weak, constant magnetic field $\delta\mathcal{H}$, applied along the z-axis. In this case a new term appears in the Hamiltonian (28), namely

$$-\frac{e}{m}\delta\mathcal{H}\sum_{\vec{p}}(a^{+}_{\vec{p}+}a_{\vec{p}+}-a^{+}_{\vec{p}-}a_{\vec{p}-}), \tag{51}$$

where $+e$ is the charge of the electron and m its mass.

The compensation equations lead to the formulae

$$F_{\delta\mathcal{H}}(\vec{p},+)=\frac{1}{2}\left[1-\frac{\vartheta(\vec{p})-\frac{e}{m}\delta\mathcal{H}}{\sqrt{\left[\vartheta(p)-\frac{e}{m}\delta\mathcal{H}\right]^2+C^2(\vec{p})}}\right],$$

$$F_{\delta\mathcal{H}}(\vec{p},-)=\frac{1}{2}\left[1-\frac{\vartheta(p)+\frac{e}{m}\delta\mathcal{H}}{\sqrt{\left[\vartheta(p)+\frac{e}{m}\delta\mathcal{H}\right]^2+C^2(p),}}\right],$$

$$C(\vec{p},+)=C(\vec{p},-)\approx C(\vec{p}). \tag{52}$$

Hence

$$S^z_{\delta\mathcal{H}} = \frac{1}{2}\frac{dn}{dE}\int_{-\omega}^{\omega}\int_{-1}^{1}\frac{\vartheta + \frac{e}{m}\delta\mathcal{H}}{\sqrt{\left(\vartheta + \frac{e}{m}\delta\mathcal{H}\right)^2 + \Psi^2 x^2}}\,d\vartheta\,dx. \tag{53}$$

On perfoming this integration, we get

$$S_{\delta\mathcal{H}} = \frac{dn}{dE}\frac{e}{m}\delta\mathcal{H}\,[1 - \exp(2/3 - 6/\varrho_1)] \tag{54}$$

For the paramagnetic susceptibility we obtain

$$\chi = \frac{2e^2}{m^2}\frac{dn}{dE}\,[1 - \exp(2/3 - 6/\varrho_1)]. \tag{55}$$

From (54) we see that in the absence of a magnetic field the total spin of a system equals zero. In the presence of a magnetic field the electron-electron interaction cannot be taken into account by means of perturbation theory since we have obtained a non-analytical dependence of (54) and (55) for $\varrho_1 \approx 0$.

In this paper we have considered formally the solution of the compensation equations as an odd function of $\vec{p}$. It is understood that from the point of view of transformation (2) this leads to a state of the system with pairs of fermions with parallel spins. In order to interpret the physical picture of this state one has to examine the thermodynamics of this anomalous state.

It is a pleasure for the author to express his gratitude to Professor N. N. Bogolubov for suggesting the problem and helpful advice. The author also wishes to thank Dr D. V. Shirkov and Dr V. V. Tolmachev for valuable discussions.

REFERENCES

Bogolubov, N. N., Tolmachev, V. V., Shirkov, D. V., „*A new method in the theory of superconductivity*", Izdat. Akad. Nauk SSSR (1958). English copy in *Fortschr. Phys.*, **6**, 605 (1958). English edition — Consultants Bureau, INC. New York (1959).

Bogolubov, N. N., *Uspekhi fiz. Nauk*, **67**, 549 (1959).

Galasiewicz, Z., *P og . theor. Phys.*, **23**, 197 (1960).

Landau, L. D., *Zh. eksper. teor. Fiz.*, **30**, 1058 (1956).

Khalatnikov, I. M., Abrikosov, A. A., *Uspekhi fiz. Nauk*, **65**, 551 (1958). English copy in *Advances in Phys.* **8**, 45 (1959).

AN ELEMENTARY DERIVATION OF THE FORMULAE FOR MULTIPOLE RADIATION

By Bronisław Średniawa

Institute of Theoretical Physics, Jagellonian University, Cracow

(Received December 31, 1959)

Expressions for the field intensities of the different electromagnetic multipole radiations produced by electric charges and currents as well as by electric and magnetic polarizations varying harmonically in time are deduced by developing the retarded potentials into series. The formulae obtained are direct generalizations of the Rubinowicz-Sommerfeld formulae for dipole and quadrupole radiations and are valid in the wave zone. The general formulae for the field intensities are also given in tensor form.

Introduction

In recent years the theory of multipole radiation has been extensively developed, particularly for the needs of nuclear physics. The results of these investigations are presented e. g. in the books of Blatt and Weisskopf (1952, chap. 8), Rose (1955) and Hamilton (1959, chap. 1), where also the literature of the subject is given. Since the methods of modern theory of multipole radiation are rather complicated I have tried to get the formulae for the field intensities of the electromagnetic field in an elementary and straightforward way by simplifying and generalizing a method of Rubinowicz and Sommerfeld (Rubinowicz and Blaton 1932, Sommerfeld 1939, chap. 1). We assume that the electromagnetic radiation is produced by electric charges and currents as well as by electric and magnetic polarizations varying harmonically in time.

We begin with the decomposition of the electromagnetic field into a „charge--dependent" part and a „magnetization-dependent" part. For the first part of the field the retarded potentials are developed into series and decomposed again into two parts representing separately electric and magnetic multipole radiations. After the expressions for the field intensities of the „charge-dependent" part have been found, the field intensities for the „magnetization-dependent" part may be found without further calculations. No explicit use of the theory of groups and spherical harmonics is made. The formulae obtained are valid in the wave zone and are direct

generalizations of the Rubinowicz-Sommerfeld formulae for dipole and quadrupole radiations. Formulae for the electric and magnetic intensities of multipole radiation are also given in tensor form and are seen to be direct generalizations of the formulae for the lowest multipole radiations.

§ 1. *Decomposition of the electromagnetic field into „charge-dependent" and „magnetization-dependent" parts*

Let the charge density $\bar{\varrho}_f$ current density $\bar{\boldsymbol{j}}_f$, electric and magnetic polarization densities $\bar{\boldsymbol{P}}$ and $\bar{\boldsymbol{M}}$ be given in a spatial region V, outside of which all these quantities vanish.

The Maxwell equations (in rationalized units) are then

$$\operatorname{rot} \boldsymbol{E} + \frac{1}{c}\dot{\boldsymbol{H}} = -\frac{1}{c}\dot{\bar{\boldsymbol{M}}}, \qquad \operatorname{div} \bar{\boldsymbol{H}} = -\operatorname{div} \bar{\boldsymbol{M}}$$

$$\operatorname{rot} \boldsymbol{H} - \frac{1}{c}\dot{\boldsymbol{E}} = \frac{1}{c}(\bar{\boldsymbol{j}}_f + \dot{\bar{\boldsymbol{P}}}), \qquad \operatorname{div} \boldsymbol{E} = \bar{\varrho}_f - \operatorname{div} \bar{\boldsymbol{P}}. \tag{1}$$

From these equations follows the law of conservation of electric charge

$$\partial_t \bar{\varrho}_f + \operatorname{div} \bar{\boldsymbol{j}}_f = 0.$$

We decompose each of the fields $\boldsymbol{E}$ and $\boldsymbol{H}$ into two parts: the „charge-dependent" part $\boldsymbol{E}_1$, $\boldsymbol{H}_1$ and the „magnetization-dependent" part $\boldsymbol{E}_2$, $\boldsymbol{H}_2$:

$$\boldsymbol{E} = \boldsymbol{E}_1 + \boldsymbol{E}_2 \qquad \boldsymbol{H} = \boldsymbol{H}_1 + \boldsymbol{H}_2 \tag{2}$$

satisfying the following equations

$$\operatorname{rot} \boldsymbol{E}_1 + \frac{1}{c}\dot{\boldsymbol{H}}_1 = 0, \qquad \operatorname{div} \boldsymbol{H}_1 = 0$$

$$\operatorname{rot} \boldsymbol{H}_1 - \frac{1}{c}\dot{\boldsymbol{E}}_1 = \bar{\boldsymbol{j}}, \qquad \operatorname{div} \boldsymbol{E}_1 = \bar{\varrho}, \tag{3}$$

where

$$\bar{\boldsymbol{j}}_f + \dot{\bar{\boldsymbol{P}}} \equiv \bar{\boldsymbol{j}}, \qquad \bar{\varrho}_f - \operatorname{div} \bar{\boldsymbol{P}} \equiv \bar{\varrho}, \tag{4}$$

and

$$\operatorname{rot} \boldsymbol{E}_2 + \frac{1}{c}\dot{\boldsymbol{H}}_2 = -\frac{1}{c}\dot{\bar{\boldsymbol{M}}}, \qquad \operatorname{div} \boldsymbol{H}_2 = -\operatorname{div} \bar{\boldsymbol{M}}$$

$$\operatorname{rot} \boldsymbol{H}_2 - \frac{1}{c}\dot{\boldsymbol{E}}_2 = 0, \qquad \operatorname{div} \boldsymbol{E}_2 = 0. \tag{5}$$

From (4) we see that the following relation is satisfied

$$\partial_t \bar{\varrho} + \operatorname{div} \bar{\boldsymbol{j}} = 0. \tag{6}$$

§ 2. *Retarded potentials and the electromagnetic field for the charge-dependent part*

The vectors $\boldsymbol{E}_1$ and $\boldsymbol{H}_1$ may be derived from the retarded potentials

$$\varphi(P,t) = \frac{1}{4\pi}\int \frac{[\bar{\varrho}]}{r_{PQ}}\,dV, \qquad \boldsymbol{a}(P,t) = \frac{1}{4\pi\, c}\int \frac{[\bar{\boldsymbol{j}}]}{r_{PQ}}\,dV, \tag{7}$$

where r_{PQ} is the distance between P and the source point Q and the expressions in brackets have to be taken at the time $t-r_{PQ}/c$. We assume that $\bar{\varrho}$, ... $\bar{\boldsymbol{M}}$ oscillate harmonically in time:

$$\bar{\varrho} = \varrho\, e^{-i\omega t} + \varrho^*\, e^{i\omega t} \quad \text{etc.} \tag{8}$$

where ϱ etc. are time independent. Introducing $k = 2\pi/\lambda = \omega/c$ we may write

$$\bar{\varrho} = \varrho\, e^{-ickt} + c\cdot c \tag{8'}$$

Now we confine ourselves to the wave zone. We choose a point O so near to Q that $r \ll R$ where $\boldsymbol{r} = \overrightarrow{OQ}$, $\boldsymbol{R} = \overrightarrow{OP}$ (and $\boldsymbol{R}/R = \boldsymbol{n}$). It is well known (see e. g. Sommerfeld 1939, Vol. II, p. 61, Formula (12)) that the retarded potentials take then the form

$$\begin{pmatrix}\varphi\\ \boldsymbol{a}\end{pmatrix} = \frac{e^{ik(R-ct)}}{4\pi\, R}\int \begin{pmatrix}\varrho\\ \boldsymbol{j}/c\end{pmatrix} e^{-ik\boldsymbol{r}\cdot\boldsymbol{n}} + c\cdot c. \tag{9}$$

Developing

$$e^{-ik\boldsymbol{r}\cdot\boldsymbol{n}} = \sum \frac{(-ik)^l}{l!}(\boldsymbol{r}\cdot\boldsymbol{n})^l$$

and denoting

$$\Phi = \frac{1}{R}\, e^{ik(R-ct)} \tag{10}$$

we get

$$\varphi = \Phi \sum_{l=0}^{\infty} \varrho\,(\boldsymbol{n}\cdot\boldsymbol{r})^l\, dV + c\cdot c. \tag{11}$$

$$\boldsymbol{a} = \frac{\Phi}{c} \sum_{l=0}^{\infty} \boldsymbol{j}\,(\boldsymbol{n}\cdot\boldsymbol{r})^l\, dV + c\cdot c. \tag{12}$$

Writing equation (6) in the form

$$\operatorname{div}\boldsymbol{j} = ick\,\varrho \tag{13}$$

we see at once that in our case the total charge vanishes:

$$\int \varrho\, dV = 0 \tag{14}$$

We shall decompose now (in a manner similar to that of Fierz 1949) each term of the series for a into a part which may be expressed with the help of ϱ and another part independent of ϱ. To this purpose we start from the obvious relation

$$\int \boldsymbol{B}\, dV = -\int \boldsymbol{r} \operatorname{div} \boldsymbol{B}\, dV \tag{15}$$

valid for each $\boldsymbol{B}$ vanishing outside the finite region V. Substituting in (15)

$$\boldsymbol{B} = (\boldsymbol{n}\cdot\boldsymbol{r})^l \boldsymbol{j}, \qquad l = 1, 2, \ldots \tag{16}$$

and taking into account (13), we get from (15) and (16)

$$-ick \int (\boldsymbol{n}\cdot\boldsymbol{r})^l \varrho\, \boldsymbol{r}\, dV = \int (\boldsymbol{n}\cdot\boldsymbol{r})^{l-1} \{(\boldsymbol{n}\cdot\boldsymbol{r})\boldsymbol{j} + l(\boldsymbol{n}\cdot\boldsymbol{j})\boldsymbol{r}\}\, dV \tag{17}$$

Now we may write the l-th term of a in the form

$$\frac{(-ik)^l}{l!}\int (\boldsymbol{n}\cdot\boldsymbol{r})^l\, dV = \frac{(-ik)^l}{l!}\left\{\frac{1}{l+1}\int (\boldsymbol{n}\cdot\boldsymbol{r})^{l-1}\,[(\boldsymbol{n}\cdot\boldsymbol{j})\boldsymbol{j} + (\boldsymbol{n}\cdot\boldsymbol{j})\boldsymbol{r}]\, dV + \right.$$

$$\left. + \frac{l}{l+1}\int (\boldsymbol{n}\cdot\boldsymbol{r})^{l-1}\,[(\boldsymbol{n}\cdot\boldsymbol{r})\boldsymbol{j} - (\boldsymbol{n}\cdot\boldsymbol{j})\boldsymbol{r}]\, dV\right\}. \tag{18}$$

Taking into account (17) and putting $\boldsymbol{n}\cdot\boldsymbol{r} = r_n$, we get finally the following series for $\boldsymbol{a}$:

$$\boldsymbol{a} = \Phi \sum_{l=0}^{\infty} \frac{(-ik)^{l+1}}{(l+1)!}\int r_n^l \varrho \boldsymbol{r}\, dV + \frac{1}{c}\Phi \sum_{l=0}^{\infty}\frac{(-ik)^l\, l}{(l+1)!}\int r_n^l (\boldsymbol{r}\times\boldsymbol{j})\times\boldsymbol{n}\, dV + c\cdot c \tag{19}$$

§ 3. *The charge-dependent field of the multipole radiation*

a. Electric 2^l-pole radiation

Let us take together the successive terms for φ and for $\boldsymbol{a}$ from (11) and the first part of (19):

$$\varphi_l^{(e)} = \Phi \frac{(-ik)^l}{l!}\int r_n^l\, dV + c\cdot c., \qquad l = 1, 2, \ldots \tag{20}$$

$$\boldsymbol{a}_l^{(e)} = \Phi \frac{(-ik)^l}{l!}\int r_n^{l-1} \varrho \boldsymbol{r}\, dV + c\cdot c. \tag{21}$$

These potentials satisfy the Lorentz condition in the wave zone. Thank to the relations

$$\boldsymbol{E}_1 = -\operatorname{grad}\varphi - \frac{1}{c}\dot{\boldsymbol{a}}, \qquad \boldsymbol{H}_1 = \operatorname{rot}\boldsymbol{a} \tag{22}$$

we get from (20) and (21) the field intensities

$$\boldsymbol{E}_1^{(e)l} = -\Phi \frac{(-ik)^{l+1}}{l!} \int r_n^{l-1} \varrho \, \boldsymbol{r}_\perp \, dV + c \cdot c. \tag{23}$$

$$\boldsymbol{H}_1^{(e)l} = -\Phi \frac{(-ik)^{l+1}}{l!} \int r_n^{l-1} \boldsymbol{n} \times \varrho \, \boldsymbol{r} \, dV + c.c. = \boldsymbol{n} \times \boldsymbol{E}_1^{(e)l} \tag{24}$$

where $\boldsymbol{r}_\perp$ is the component of $\boldsymbol{r}$ perpendicular to $\boldsymbol{n}$.

These fields satisfy the symmetry relations

$$\boldsymbol{E}_1^{(e)l}(-\boldsymbol{R}) = (-1)^{l-1} \boldsymbol{E}_1^{(e)l}(\boldsymbol{R}), \quad \boldsymbol{H}_1^{(e)l}(-\boldsymbol{R}) = (-1)^l \boldsymbol{H}_1^{(e)l}(\boldsymbol{R}) \tag{25}$$

characterizing the electric multipole radiation.

Taking into account (4), we can finally write (23) and (24) in the form

$$\boldsymbol{E}_1^{(e)l} = -\Phi \frac{(-ik)^{l+1}}{l!} \int r_n^{l-1} (\varrho_f - \operatorname{div} \boldsymbol{P}) \, \boldsymbol{r}_\perp \, dV + c.c. \tag{26}$$

$$\boldsymbol{H}_1^{(e)l} = -\Phi \frac{(-ik)^{l+1}}{l!} \int r_n^{l-1} \boldsymbol{n} \times \boldsymbol{r} \, (\varrho_f - \operatorname{div} \boldsymbol{P}) \, dV + c.c. \tag{27}$$

b. Magnetic 2^l-pole radiation

The remaining part of $\boldsymbol{a}$ in (19) yields

$$\varphi_l^{(m)} = 0$$

$$\boldsymbol{a}_l^{(m)} = \frac{1}{c} \Phi \frac{(-ik)^l \, l}{(l+1)!} \int r_n^{l-1} (\boldsymbol{r} \times \boldsymbol{j}) \times \boldsymbol{n} \, dV + c.c., \quad l = 1, 2, \ldots \tag{28}$$

the Lorentz condition for $\varphi_l^{(m)}$ and $\boldsymbol{a}_l^{(m)}$ is again satisfied and simple calculations lead from (3) to the following expression for the field intensities

$$\boldsymbol{E}_1^{(m)l} = \Phi \frac{(-ik)^{l+1} \, l}{(l+1)!} \int r_n^{l-1} \boldsymbol{n} \times \{\boldsymbol{r} \times (\boldsymbol{j}_f - ick \boldsymbol{P})\} \, dV + c.c. \tag{29}$$

$$\boldsymbol{H}_1^{(m)l} = -\Phi \frac{(-ik)^{l+1} \, l}{(l+1)! \, c} \int r_n^{l-1} \{\boldsymbol{r} \times (\boldsymbol{j}_f - ick \boldsymbol{P})\}_\perp \, dV + c.c. = \boldsymbol{n} \times \boldsymbol{E}_1^{(m)l} \tag{30}$$

which satisfy the symmetry relations

$$\boldsymbol{E}_1^{(m)l}(-\boldsymbol{R}) = (-1)^l \boldsymbol{E}_1^{(m)l}(\boldsymbol{R}), \quad \boldsymbol{H}_1^{(m)l}(-\boldsymbol{R}) = (-1)^{l+1} \boldsymbol{H}_1^{(m)l}(\boldsymbol{R}) \tag{31}$$

characterising the magnetic multipole radiation.

c. Analogies

Comparing formuale (23) and (24) on the one hand and (29) and (30) on the other hand, we see that the replacing in (23) and (24)

$$\text{of } \boldsymbol{E}_1^l, \quad \boldsymbol{H}_1^l, \quad \varrho \boldsymbol{r}$$

$$\text{by } \boldsymbol{H}_1^l, \quad -\boldsymbol{E}_1^l, \quad \frac{l}{(l+1)\,c}\,\boldsymbol{r}\times\boldsymbol{j}$$

gives exactly formulae (27) and (28) for all multipole radiations.

§ 4. The magnetization-dependent part of the multipole radiation

Upon replacing

$$\boldsymbol{E}_2, \quad \boldsymbol{H}_2, \quad \dot{\bar{\boldsymbol{M}}}, \quad -\operatorname{div} \bar{\boldsymbol{M}}$$

by (32)

$$-\boldsymbol{H}_1, \quad \boldsymbol{E}_1, \quad -\bar{\boldsymbol{j}}, \quad \bar{\varrho}$$

equations (5) change into (3) and we can write directly the expressions for the magnetization-dependent part of the multipole radiation without further calculations.

a. Electric multipole radiation

From (29) and (30) we get

$$\boldsymbol{E}_2^{(e)l} = -\Phi\,\frac{(-ik)^{l+2}\,l}{(l+1)!}\int r_n^{l-1}\,(\boldsymbol{r}\times\boldsymbol{M})_\perp\,dV \tag{33}$$

$$\boldsymbol{H}_2^{(e)l} = -\Phi\,\frac{(-ik)^{l+2}\,l}{(l+1)!}\int r_n^{l-1}\,\boldsymbol{n}\times(\boldsymbol{r}\times\boldsymbol{M})\,dV = \boldsymbol{n}\times\boldsymbol{E}_2^{(e)l} \tag{34}$$

It is easily seen that these vectors have the same symmetry properties as the fields (23) and (24), and therefore they belong to the electric multipole radiation.

b. Magnetic multipole radiation

On the other hand from (24) and (25) we get the field intensities

$$\boldsymbol{E}_2^{(m)l} = -\Phi\,\frac{(-ik)^{l+1}}{l!}\int r_n^{l-1}\,\boldsymbol{n}\times\boldsymbol{r}\operatorname{div}\boldsymbol{M}\,dV \tag{35}$$

$$\boldsymbol{H}_2^{(m)l} = \Phi\,\frac{(-ik)^{l+1}}{l!}\int r_n^{l-1}\,\boldsymbol{r}_\perp \operatorname{div}\boldsymbol{M}\,dV \tag{36}$$

possessing the symmetry of the magnetic multipole radiation.

c. Total field

We obtain the total electric or magnetic 2^l-pole radiation by superposition

$$\begin{aligned} \boldsymbol{E}^{(e)l} &= \boldsymbol{E}_1^{(e)l} + \boldsymbol{E}_2^{(e)l}, \qquad \boldsymbol{H}^{(e)l} = \boldsymbol{H}_1^{(e)l} + \boldsymbol{H}_2^{(e)l} \\ \boldsymbol{E}^{(m)l} &= \boldsymbol{E}_1^{(m)l} + \boldsymbol{E}_2^{(m)l}, \qquad \boldsymbol{H}^{(m)l} = \boldsymbol{H}_1^{(m)l} + \boldsymbol{H}_2^{(m)l} \end{aligned} \tag{37}$$

From these formulae the energy and the Pointing vector for different multipole radiations may be easily calculated.

§ 5. *Multipole moments. Tensor form of the formulae for multipole radiation*

We define the time-independent multipole moments as follows:

a. Electric multipole moments

$$\begin{aligned} Q^1_{i_1 \dots i_l} &= \int x_{i_1 \dots} x_{i_l} (\varrho_f - \partial_i P_i)\, dV \\ Q^2_{i_1 \dots i_{l+1}} &= \int x_{i_1 \dots} x_{i_{l-1}} (x_{i_l} M_{i_{l+1}} - x_{i_{l+1}} M_{i_l})\, dV \end{aligned} \tag{38}$$

b. Magnetic multipole moments

$$\begin{aligned} J^1_{i_1 \dots i_{l+1}} &= \int x_{i_1 \dots} x_{i_{l-1}} \{x_{i_l} (j_{i_{l+1}} - ick\, P_{i_{l+1}}) - x_{i_{l+1}} (j_{i_l} - ick\, P_{i_l})\}\, dV \\ J^2_{i_1 \dots i_l} &= \int x_{i_1 \dots} x_{i_l} \partial_i M_i\, dV \end{aligned} \tag{39}$$

(All indices run form 1 to 3). These multipole moments are tensors with respect to linear spatial transformations. The multipole moments depending on time are then

$$Q^1_{i_1 \dots i_l} = e^{-ickt} Q_{i_1 \dots i_l} + c \cdot c. \quad \text{etc.} \tag{40}$$

Introducing the completely antisymmetrical tensor ε_{ikl}, taking account of the space--time dependence of Φ, and putting

$$\begin{aligned} A_{i_1 \dots i_l} &= \frac{n_{i_1} \dots n_{i_l}}{4\pi\, Rl!\, c^{l+1}}, \\ \overline{Q}^{1(l+1)}_{i_1 \dots i_l} &= \frac{d^{l+1}}{dt^{l+1}} \overline{Q}^1_{i_1 \dots i_l} \end{aligned}$$

we see that we can write the formulae for the electric and magnetic multipole radiation in tensor form as follows:

a. Electric 2^l-pole radiation (from (26), (33) and (27), (34)):

$$E_i^{(e)l} = A_{i_1 \dots i_l} \left\{ 2\,\overline{Q}^{1(l+1)}_{i_1 \dots i_{l-1}} [_{i_l} n_{i]} - \frac{l}{(l+1)\,c}\,\overline{Q}^{2(l+2)}_{i_1 \dots i_{l+1}}\,\varepsilon_{i i_{l+1} k}\, n_k \right\}_{t - \frac{R}{c}}$$

$$H_i^{(e)l} = A_{i_1 \dots i_l} \left\{ \overline{Q}^{1(l+1)}_{i_1 \dots i_{l-1} k}\,\varepsilon_{iki_1} - \frac{l}{(l+1)\,c}\,\overline{Q}^{2(l+2)}_{i_1 \dots i_l i} \right\}_{t - \frac{R}{c}}$$

b. Magnetic 2^l — pole radiation (from (29), (35) and (30), (36))

$$E_i^{(m)l} = A_{i_1 \dots i_l} \left\{ \overline{J}^{2(l+1)}_{i_1 \dots i_{l-1} k}\,\varepsilon_{iki_l} - \frac{l}{(l+1)\,c}\,\overline{J}^{1(l+1)}_{i_1 \dots i_{l-1} i} \right\}_{t - \frac{R}{c}}$$

$$H_i^{(m)l} = A_{i_1 \dots i_l} \left\{ \frac{l}{(l+1)\,c}\,\overline{J}^{1(l+1)}_{i_1 \dots i_{l+1}}\,\varepsilon_{i i_{l+1} k}\, n_k - 2\,\overline{J}^{2(l+1)}_{i_1 \dots i_{l-1} [i_l}\, n_{i]} \right\}_{t - \frac{R}{c}}$$

These formulae are the generalizations of the well known formulae for the lowest multipole radiations.

§ 6. *Final remarks*

a. We may get rid of div $\boldsymbol{P}$ and div $\boldsymbol{M}$ in formulae (26) and (27) by application of the identity

$$(\nabla \cdot \boldsymbol{b})\,\boldsymbol{a} = \boldsymbol{a}\,\operatorname{div}\boldsymbol{b} + (\boldsymbol{b} \cdot \nabla)\,\boldsymbol{a}$$

for $\boldsymbol{a} = r_n^{l-1}\,\boldsymbol{r}_\perp$ or $\boldsymbol{a} = r_n^{l-1}\,\boldsymbol{n} \times \boldsymbol{r}$ and $\boldsymbol{b} = \boldsymbol{P}$ or $\boldsymbol{b} = \boldsymbol{M}$.

We get then after elementary calculations instead of (26) and (27)

$$\boldsymbol{E}_1^{(e)l} = -\,\Phi\,\frac{(-\,ik)^{l+1}}{l!} \left\{ \int r_n^{l-1} (\varrho_f \boldsymbol{r}_\perp + \boldsymbol{P}_\perp)\,dV + (l-1) \int r_n^{l-2}\,P_n\,\boldsymbol{r}_\perp\,dV \right.$$

$$\boldsymbol{H}_1^{(el)} = \boldsymbol{n} \times \boldsymbol{E}_1^{(e)l}$$

and instead of (35) and (36)

$$E_2^{(m)l} = -\,\Phi\,\frac{(-\,ik)^{l+1}}{l!} \left\{ \int r_n^{l-1} \boldsymbol{M} \times \boldsymbol{n}\,dV - (l-1) \int r_n^{l-2}\,M_n\,\boldsymbol{n} \times \boldsymbol{r}\,dV \right\}$$

$$H_2^{(m)l} = \boldsymbol{n} \times E_2^{(m)l}$$

b. It is easily seen that we may express the retarded potentials in a way analogous to the static potential ($R = X,\ Y,\ Z$)

$$\varphi_l^{(e)} = \frac{1}{R}\,\frac{(-1)^l}{l!}\,Q^1_{i_1 \dots i_l}\,\frac{\partial^l}{\partial X_{i_1} \dots \partial X_{i_l}}\,e^{ik(R-ct)}$$

and similarly for $a_l^{(e)}$, $a_l^{(m)}$ differentiating the "Maxwell R" (in the exponent) instead of the "Coulomb R" (in the denominator).

I wish to express my thanks to Professor Jan Weyssenhoff and to Miss A. Kowalska for interesting discussions and remarks.

REFERENCES

Blatt, J. M. and Weisskopf, V. F., *Theoretical Nuclear Physics*, Wiley, N. York, London 1952.
Fierz, M., *Helv. phys. Acta*, **22**, 489 (1949).
Hamilton, J., *The theory of elementary particles*, Clarendon Press, Oxford 1959.
Rose, M. E., *Multipole fields*, Wiley, N. York 1955.
Rubinowicz, A. and Blaton, J., *Erg. der exacten Naturwiss.* **11**, 176 (1932).
Sommerfeld, A., *Atombau und Spektrallinien*, Vieweg, Berlin 1939.

EINFLUSS DER KRISTALLGITTERBEGRENZUNGEN AUF DIE SPINWELLENRESONANZ IN EINEM FERRIMAGNETIKUM

Von A. R. Ferchmin

Institut für Theoretische Physik der Adam Mickiewicz Uniwersität, Poznań

(Eingegangen am 13 Januar 1960)

Es wird die Spinwellenresonanz untersucht, welche von dem homogenen magnetischen Wechselfeld in einer dünnen ferrimagnetischen Schicht erzeugt wird. Man erhält eine Lösung von N halbklassischen Spinbewegungsgleichungen. Die Form dieser Lösung hängt davon ab, ob an beiden Enden einer linearen ferrimagnetischen Kette die Spinmomenten gleich oder ungleich sind. Es wird gezeigt, dass man ausser den gewöhnlichen Spinwellen mit einer periodischen Amplitude auch die aperiodische Spinwellen erregen kann.

I. Einleitung

Die den Kristall begrenzenden Flächen verursachen Spinwellenreflexionen und unter bestimmten Bedingungen können stehende Spinwellen entstehen. Die Spins an der Oberfläche befinden sich unter der Wirkung anderer Kräfte als die im Innern des Kristalls, und zwar wegen der fehlenden Nachbarn. Dieser Gedanke wurde zuerst von Kittel (1958) ausgesprochen, und er hat ihn zum Voraussagen der sogenannten Spinwellenresonanz im homogenen Hf Feld geführt. Zwischen dem halbklassischen Bild der Spinwellen und den mechanischen Wellen in Kristallen besteht eine Analogie. Das Problem der Wellenausbreitung in endlichen Kristallgittern ist schon lange bekannt (siehe z.B. Born and Huang 1954). Man sieht in endlichen Gittern von der sogenanntem Periodizitätsbedingung ab und setzt andere Randbedingungen voraus. Der Fall einer endlichen linearen Kette wurde auch von Wallis (1957) diskutiert, der u.a. einige Sonderlösungen, die er Oberflächenschwingungen nannte, gefunden hat. Kittel (1958) bediente sich der Bewegungsgleichung für den Spin in einer Form, die für stetige Medien gilt, d.h. wenn die Gitterkonstante a nach Null oder die Wellenlänge nach Unendlichkeit strebt. Die Behandlung eines Kristalls wie ein stetiges Medium, was bei grossen Kristallen begründet ist, kann in dünnen Schichten manche Diskontinuitäts- und Oberflächeneffekte (z. B. ungleiche Differenzen der sukzessiven Wellenzahlen der stehenden Wellen) verschleiern. Kittel nimmt ein Ferromagnetikum in Betracht; Orbach und Pincus (1959) und unabhängig von ihnen Cofta (1960a) haben den Fall eines antiferromagnetischen Stoffes behandelt.

Der allgemeine Fall eines ferrimagnetischen Stoffes, in welchem die anderen Fälle mitenthalten sind, wurde von Cofta (1960b) diskutiert. Cofta nahm die von Kittel (1958) formulierten Randbedingungen an, und löste dann nur die Bewegungsgleichungen für die inneren Spins. Auf diese Weise nahm er nur indirekt den Einfluss der Kristallenden in Betracht. Eine solche Betrachtungsweise ist aber vom mathematischen Standpunkt nicht begründet, da die Spinbewegungsgleichungen, wie wir weiter sehen werden, miteinander gekoppelt sind, so dass sie ein System von Gleichungen bilden. (Siehe auch Fox 1957 S. 5) Dieses System kann nur dann in einzelne Gleichungen aufgespaltet werden, wenn alle Gleichungen (bzw. Gleichungspaare) dieselbe Form haben, d. h. wenn man von den Randgleichungen absehen kann. Für sehr dünne Schichten kann aber eine solche Annahme eine ungenügende Annäherung sein.

Wir wollen hier die oben erwähnten Annäherungen vermeiden und eine ferrimagnetische endliche lineare Kette untersuchen. Es wird auch der spezielle Fall einer ferromagnetischen Kette untersucht.

II. Spinwellen in einer endlichen Kette

A. Ungleiche Spins an den Enden

Betrachten wir eine lineare Kette in der Richtung der z-Achse von äquidistant fixierten Spins von zwei Arten: $\boldsymbol{S}_\alpha$ und $\boldsymbol{S}_\beta$, die alternierend gesetzt sind, so dass ein Atom mit dem Spinmoment $\boldsymbol{S}_\alpha$ zwei $\boldsymbol{S}_\beta$ als nächste Nachbarn hat und umgekehrt. Die Bewegung eines Spins wird durch die bekannte halbklassische Gleichung

$$\frac{\partial \boldsymbol{S}_l}{\partial t} = \frac{2J}{\hbar}\, \boldsymbol{S}_l \times \sum_r \boldsymbol{S}_{l+r} + \gamma\, \boldsymbol{S}_l \times \boldsymbol{H} \tag{1}$$

beschrieben, wo Σ die Summe der Nachbarspins und J den Austauschintegral zwischen Nachbarspins bedeutet. Die Spins an den Enden der Kette können ungleich oder gleich sein. Wenn sowohl die Zahl der $\boldsymbol{S}_\alpha$ Spins, als auch die Zahl der $\boldsymbol{S}_\beta$ Spins N beträgt, so ist der erste Fall realisiert. Die Spins befinden sich in einem zur z-Achse parallelen, statischen magnetischen Feld $\boldsymbol{H}_o$. Dieses Feld enthält auch das effektive Anisotropiefeld, das auf die inneren Spins wirkt und darum unterscheiden wir $\boldsymbol{H}_{\alpha 0}$ und $\boldsymbol{H}_{\beta 0}$. Wir folgen Kittel (1958) und nehmen an, dass in der z-Richtung das Oberflächenanisotropiefeld $\boldsymbol{H}_{a\alpha}$ auf den Spin $\boldsymbol{S}_{\alpha 1}$, und $\boldsymbol{H}_{a\beta}$ auf den $\boldsymbol{S}_{\beta 2N}$ wirkt. Wenn man annimmt, dass $S_\alpha^z = S_\alpha = \text{const.}$ und $S_\beta^z = S_\beta = \text{const.}$, findet man mit Hilfe der Formel (1) für die y-Komponenten der Spins:

$$\frac{\partial S_{\alpha 1}^y}{\partial t} = -\frac{2J}{\hbar}\,(S_{\alpha 1}^x\, S_\beta - S_{\beta 2}^x\, S_\alpha) - \gamma S_{\alpha 1}^x\,(H_{\alpha 0} + H_{a\alpha})$$

. . . . . .

$$\frac{\partial S_{\alpha\, 2j-1}^y}{\partial t} = -\frac{2J}{\hbar}\,[2\, S_{\alpha\, 2j-1}^x\, S_\beta - (S_{\beta\, 2j-2}^x + S_{\beta\, 2j}^x)\, S_\alpha] - \gamma\, S_{\alpha\, 2j-1}^x\, H_{\alpha 0}$$

$$\frac{\partial S^y_{\beta\,2j}}{\partial t} = -\frac{2J}{\hbar}\left[2\,S^x_{\beta\,2j}\,S_\alpha - (S^x_{\alpha\,2j-1} + S^x_{\alpha\,2j+1})\,S_\beta\right] - \gamma\,S^x_{\beta\,2j}\,H_{\beta 0}$$

. . . . . .

$$\frac{\partial S^y_{\beta\,2N}}{\partial t} = -\frac{2J}{\hbar}\,(S^x_{\beta\,2N}\,S_\alpha - S^x_{\alpha\,2N-1}S_\beta) - \gamma\,S^x_{\beta\,2N}\,(H_{\beta 0} + H_{a\beta}) \qquad (2)$$

und ähnliche Systeme von $2N$ Gleichungen für $\dot{S}^x$ und $\dot{S}^z$.

Setzen wir

$$\begin{aligned} S^x_{\alpha\,2j-1} &= A_{2j-1}\,e^{i\omega t},\\ S^x_{\beta\,2j} &= A_{2j}\,e^{i\omega t},\\ S^y_{\alpha\,2j-1} &= i\,A_{2j-1}\,e^{i\omega t},\\ S^y_{\beta\,2j} &= i\,A_{2j}\,e^{i\omega t}, \end{aligned} \qquad (3)$$

so können wir schreiben:

$$\begin{aligned} &(u+a)\,A_1 + A_2 = 0\\ &\quad\ldots\ldots\ldots\\ &A_{2j-2} + u\,A_{2j-1} + A_{2j} = 0\\ &A_{2j-1} + v\,A_{2j} + A_{2j+1} = 0\\ &\quad\ldots\ldots\ldots\\ &A_{2N-1} + (v+b)\,A_{2N} = 0, \end{aligned} \qquad (4)$$

($2N$ Gleichungen)

wo

$$\begin{aligned} &u = (\omega - 2\omega_{e\beta} + \omega_{\alpha 0})/\omega_{e\alpha},\ v = (\omega - 2\omega_{e\alpha} - \omega_{\beta 0})/\omega_{e\beta},\\ &a = (\omega_{e\beta} - \omega_{a\alpha})/\omega_{e\alpha},\ b = (\omega_{e\alpha} - \omega_{a\beta})/\omega_{e\beta},\\ &\omega_{e\alpha} = (2J/\hbar)\,S_\alpha,\ \omega_{e\beta} = (2J/\hbar)\,S_\beta,\\ &\omega_{\alpha 0} = \gamma\,H_{\alpha 0},\ \omega_{\beta 0} = \gamma\,H_{\beta 0},\ \omega_\alpha = \gamma\,(H_{\alpha 0} + H_{a\alpha}),\ \omega_\beta = \gamma\,(H_{\beta 0} + H_{a\beta}),\\ &\omega_{a\alpha} = \omega_\alpha - \omega_{\alpha 0},\ \omega_{a\beta} = \omega_\beta - \omega_{\beta 0}. \end{aligned} \qquad (5)$$

Das System von Gleichungen für die S^x—Komponenten hat dieselbe Form (4) und das für die S^z—Komponenten geltende System ist identisch erfüllt von den Lösungen (3) mit beliebigen Amplituden A.

Die säkulare Gleichung für das Gleichungssystem (4) ist:

$$D_{2N}(a, b) = 0, \qquad (6)$$

wo

$$D_{2N}(a,b) = \begin{vmatrix} u+a & 1 & & & & & \\ 1 & v & 1 & & & & \\ & \cdot & \cdot & \cdot & & & \\ & & \cdot & \cdot & \cdot & & \\ & & & 1 & u & 1 & \\ & & & & 1 & v+b & \end{vmatrix}_{2N} \tag{7}$$

($2N$ markt an die Ordnung der Determinante).
Wir folgen jetzt Wallis (1957) und Rutherford (1947; 1951) und führen einen Parameter Θ ein, indem wir

$$(uv)^{\frac{1}{2}} = -2\cos\Theta \tag{8}$$

setzen. Mit Hilfe der Rechenmethoden von Rutherford kann die Determinante (7) nun in der Form

$$D_{2N}(a,b) = \frac{uv+bu+av}{-(uv)^{\frac{1}{2}}\sin\Theta}\sin 2N\Theta + (ab-1)\frac{\sin(2N-1)\Theta}{\sin\Theta} \tag{9}$$

ausgedrückt werden.

Betrachten wir zuerst einen Sonderfall, wo die Oberflächenanisotropie klein ist, d.h. $\omega_a \ll \omega_e$. Gerade in dieser Annäherung hat Kittel (1958) den ferromagnetischen Fall untersucht. Es folgt dann (s. Ausdruck 5) $a = b^{-1}$, $ab = 1$ und die säkulare Gleichung (6) wird in die Form

$$(uv+bu+av)\frac{\sin 2N\Theta}{\sin 2\Theta} = 0 \tag{10}$$

übergeführt.

Die Lösungen dieser vereinfachten Gleichung findet man aus der Beziehung:

$$\frac{\sin 2N\Theta}{\sin 2\Theta} = 0 \tag{11}$$

oder

$$uv+bu+av = 0. \tag{12}$$

Die Lösungen von (11) sind:

$$\Theta_n = n(\pi/2N), \quad 1 \leqslant n \leqslant N-1, \tag{13}$$

wo n eine ganze Zahl ist. Aus der Formel (8) folgt mit (5) (ohne Annäherung $\omega_a \ll \omega_e$)

$$\omega = \sigma\left[1 \pm \frac{\omega_{e\alpha} - \omega_{e\beta} - \frac{1}{2}(\omega_{\alpha 0} - \omega_{\beta 0})}{\sigma\varepsilon}(\cos^2\Theta + \varepsilon^2\sin^2\Theta)^{\frac{1}{2}}\right], \tag{14}$$

wo

$$\varepsilon^2 = \frac{[\omega_{e\alpha} - \omega_{e\beta} - \frac{1}{2}(\omega_{\alpha 0} - \omega_{\beta 0})]^2}{[\omega_{e\alpha} + \omega_{e\beta} - \frac{1}{2}(\omega_{\alpha 0} - \omega_{\beta 0})]^2 + 2\omega_{e\beta}(\omega_{\alpha 0} - \omega_{\beta 0})},$$

$$\sigma = \omega_{e\alpha} + \omega_{e\beta} + \tfrac{1}{2}(\omega_{\alpha 0} + \omega_{\beta 0}).$$

Die Vorzeichen $\pm$ gelten für zwei Freqenzzweige, welche wir wegen der Analogie zu den elastischen Wellen in zweikomponentigen Kristallen, akustischen und optischen Zweig nennen werden. Diese Formel (14), die sich auch in der Weise

$$\omega = \sigma \pm \{[\omega_{e\alpha} - \omega_{e\beta} - \tfrac{1}{2}(\omega_{\alpha 0} - \omega_{\beta 0})]^2 + 4\omega_{e\alpha}\omega_{e\beta}\cos^2\Theta\}^{\frac{1}{2}} \tag{15}$$

schreiben lässt, stimmt mit der Formel (9) von Cofta (1960b) überein, wenn wir den Parameter Θ durch $k_p a$ (k_p — Wellenzahl) ersetzen. Der Ausdruck (14) gibt uns also eine physikalische Deutung des Parameters Θ: er ist dem Spinwellenimpuls proportional. Wir werden ihn ferner kurz „Impuls" nennen. Der Ausdruck (14) stellt also eine Dispersionsformel dar. Wenn wir in (15) $S_\alpha = S_\beta$ bzw. $S_\alpha = -S_\beta$ setzen, so bekommen wir Formeln, die entsprechend für Ferromagnetika oder Antiferromagnetika gültig sind. Eine ausführliche Analyse der Formel (15) befindet sich in der Arbeit von Cofta (1960b).

Als mögliche Lösungen von (12) findet man leicht:

$$\omega = \omega_0, \tag{16}$$

$$\omega = \sigma_0 = \omega_{e\alpha} + \omega_{e\beta} + \omega_0. \tag{17}$$

(In der benutzten Näherung können wir $\omega_{\alpha 0} = \omega_{\beta 0} = \omega_0$ setzen). Wenn wir in (14) die $N-1$ Werte der Θ von (13) einsetzen, bekommen wir $2N-2$ Werte der ω, welche mit zwei Lösungen (16) und (17) die $2N$ Werte der Frequenz bilden. Die Lösung (17) liegt in dem „verbotenen" Bereich zwischen dem akustischen und optischen Zweig der Frequenz. Der dieser Lösung entsprechende Impuls ist komplex und wird durch die Formel

$$\Theta = \tfrac{1}{2}\pi + i \operatorname{ar\,sin\,hyp}(q/2) \tag{18}$$

ausgedrückt, wo

$$q = (S_\alpha - S_\beta)(S_\alpha S_\beta)^{-\frac{1}{2}}$$

Wenn wir die Oberflächenanisotropie nicht vernachlässigen können, lässt sich die säkulare Gleichung nur graphisch lösen. In diesem Falle bekommen wir $2N$ Werte für Θ. Sie sind nicht ganz genau gleich voneinander entfernt, wie diejenigen im Ausdruck (13). Falls die Oberflächenanisotropie, wenn auch nicht vernachlässigbar, aber klein genug ist, bekommt man u.a. auch eine komplexe Lösung.

B. Gleiche Spins an den Enden

Untersuchen wir jetzt eine lineare Kette mit gleichen S_α — Spins an ihren Enden. N sei eine gerade Zahl. Die Kette besteht aus $N+1$ S_α—Spins und N S_β—Spins. Die Bewegungsgleichungen für die inneren Spins sind in (2) gegeben. Die Bewegungsgleichungen für die Spins an den Enden haben aber die Form:

$$\frac{\partial S^y_{\alpha 1}}{\partial t} = -\frac{2J}{\hbar}(S^x_{\alpha 1} S_\beta - S^x_{\beta 2} S_\alpha) - \gamma S^x_{\alpha 1}(H_{\alpha 0} + H_{a\alpha}),$$

$$\frac{\partial S^y_{\alpha\, 2N+1}}{\partial t} = -\frac{2J}{\hbar}(S^x_{\alpha\, 2N+1} S_\beta - S^x_{\beta\, 2N} S_\alpha) - \gamma S^x_{\alpha\, 2N+1}(H_{\alpha 0} + H_{a\alpha}). \tag{19}$$

Setzen wir in diese Differentialgleichungen die Lösungen (3) ein. Unter Verwendung von (5) können wir schreiben:

$$\begin{aligned}
&(u + a) A_1 + A_2 = 0\\
&\cdots\cdots\cdots\cdots\\
&A_{2j-2} + uA_{2j-1} + A_{2j} = 0\\
&A_{2j-1} + vA_{2j} + A_{2j+1} = 0\\
&\cdots\cdots\cdots\cdots\\
&A_{2N} + (u + a)A_{2N+1} = 0
\end{aligned} \tag{20}$$

($2N+1$ Gleichungen)

Wir führen nun, Wallis (1957) folgend, die symmetrischen Präzessionsradien

$$\begin{aligned}
&A^+_{\alpha j} = \tfrac{1}{2}[A_{2j-1} + A_{2(N-j+2)-1}],\\
&A^+_{\beta j} = \tfrac{1}{2}[A_{2j} + A_{2(N-j+1)}]\\
&\text{mit } 1 \leqslant j \leqslant (N/2),
\end{aligned} \tag{21a}$$

$$\begin{aligned}
&A^+_{\alpha j} = A_{2j-1}\\
&\text{mit } j = (N/2) + 1
\end{aligned} \tag{21b}$$

und die antisymmetrischen Präzessionsradien

$$\begin{aligned}
&A^-_{\alpha j} = \tfrac{1}{2}[A_{2j-1} - A_{2(N-j+2)-1}],\\
&A^-_{\beta j} = \tfrac{1}{2}[A_{2j} - A_{2(N-j+1)}]
\end{aligned} \tag{22}$$

mit $1 \leqslant j \leqslant (N/2)$ ein. Die A^+ erfüllen ein System von $N+1$ linearen homogenen algebraischen Gleichungen. Die A^- erfüllen ein anderes System von N homogenen Gleichungen. Beide säkulare Gleichungen für die symmetrischen und antisymetrischen Lösungen A^+ und A^- lassen sich in der Form:

$$(u/v)^{\frac{1}{2}} \cos(N + 1)\,\Theta - a \cos N\Theta = 0 \tag{23}$$

(für symmetrische Präzessionsradien) bzw.

$$\frac{\sin(N+1)\,\Theta}{\sin\Theta} - a\,(v/u)^{\frac{1}{2}}\,\frac{\sin N\Theta}{\sin\Theta} = 0 \tag{24}$$

(für antisymmetrische Präzessionsradien) schreiben. Aus (23) und (24) erhält man mittels graphischen Methoden $2N+1$ Werte von Θ, und zwar $N+1$ Werte von Θ im Bereich $0 \leqslant \Theta \leqslant \pi/2$ (akustischer Zweig), und N Werte von Θ im Bereich $(\pi/2) \leqslant \Theta \leqslant \pi$ (optischer Zweig). Man bediente sich hier ebenso wie Wallis (1957) des sogenannten „Schemas ausgestreckter Zonen". Für kleine Θ aus dem akustischen Zweig gilt annähernd die Formel (13). Diese Annäherung ist um so besser, je kleiner Θ_n und je länger die Kette ist.

Für die symmetrische Wellen erhält man:

$$A_{2j} = C\,(-1)^{2j-1}\,\frac{\cos(N-2j+1)\,\Theta}{\cos N\Theta},$$

$$A_{2j-1} = C\,(-1)^{2j-1}\,(u/v)^{\frac{1}{2}}\,\frac{\cos(N-2j+2)\,\Theta}{\cos N\Theta}. \tag{25}$$

Es wird später bewiesen, dass die antisymmetrische Wellen durch das äussere Wechselfeld nicht erregt werden. Deshalb brauchen wir sie nicht berücksichtigen. Die Form dieser Lösungen zeigt uns noch einmal, dass Θ der Wellenzahl proportional ist.

Falls $S_\alpha, S_\beta > 0$, d.h. die Spinmomente gleichgerichtet sind, und die Kette genügend lang ist, haben wir nur N reelle Wertee von Θ im akustischen Zweig und $N-1$ reelle Werte von Θ im optischen Zweig. Die zwei fehlenden Θ-Werte sind komplex. Durch Substitution von $\Theta = (\pi/2)+i\psi$ in die Gleichung (23) für die symmetrischen Wellen, und $\Theta = (\pi/2)-i\psi$ in die Gleichung (24) für die antisymmetrischen Wellen erhält man Gleichungen, welche man graphisch lösen kann und auf diese Weise findet man jene zwei komplexen Werte von Θ. Die entsprechenden Werte von ω (sie sind immer reell) liegen zwischen dem akustischen und dem optischen Zweig. Wir erhalten dann gedämpfte Wellen mit einer Amplitude, welche von den Enden der Kette zu ihrer Mitte abfällt. Wir werden sie „Oberflächenspinwellen" nennen.

III. Sonderfall: eine ferromagnetische Kette

In diesem Falle haben wir $S_\alpha = S_\beta = S$, und $(u/v) = 1$. Die Wellenamplituden (25) haben also eine vereinfachte Form. Wir wollen sie mit den Lösungen Kittels (1958) vergleichen. In (25) haben wir aber den Anfangspunkt der Koordinaten in der Mitte der Kette, während er sich bei Kittel auf der Oberfläche befindet. Darum müssen wir den Cosinus anders ausdrücken; nämlich:

$$\cos(N-2j+1)\,\Theta/\cos N\Theta = \operatorname{tg} N\Theta \sin(2j-1)\,\Theta + \cos(2j-1)\,\Theta. \tag{26}$$

Wenn $\Theta_n \cong n(\pi/2N)$, dann ist $\operatorname{tg} N\Theta$ für ungerade n sehr gross. In diesem Fall also spielen nur die Sinuswellen mit ungeraden n eine Rolle. Das stimmt mit der Behauptung

Kittels überein. Aus einer Analyse der graphisch ermittelten Werte von Θ ergibt sich, dass die Annäherung (13) um so weniger genau ist, je grösser das n und je kleiner das N ist. Deshalb können in sehr dünnen Schichten bei Resonanzen höherer Ordnung auch Cosinuswellen und Sinuswellen von geradem n erscheinen[1].

IV. Erregung der Spinwellen

Nimmt man an, dass ausser dem statischen magnetischen Feld $\boldsymbol{H}_0$ auf die Spins noch ein äusseres, zirkular polarisiertes Feld in einer zur z-Achse senkrechten Richtung einwirkt, so betragen die Komponenten des totalen Feldwektors $\boldsymbol{H}$, welche auf einen S_α-Spin wirken:

$$H_x = h_0 \exp(i\omega t),\quad H_y = ih_0 \exp(i\omega t),\quad H_z = H_{\alpha 0}\ (\text{bzw. } H_z = H_{\alpha 0} + H_{a\alpha})$$

Einen analogen Ausdruck erhalten wir für einen S_β-Spin. Nehmen wir zuerst eine Kette von $2N$ Atomen in Betracht. Die Bewegungsgleichungen für die Spins sind in diesem Falle nicht homogen. Wenn wir als Lösung den Ausdruck (3) mit

$$A_{2j-1} = A'_{2j-1} + B_\alpha$$

und

$$A_{2j} = A'_{2j} + B_\beta \tag{27}$$

ansetzen, finden wir

$$\begin{aligned}
&(u + a)\, A'_1 + A'_2 = C_\alpha \\
&\cdots\cdots\cdots\cdots \\
&A'_{2j-2} + u\, A'_{2j-1} + A'_{2j} = 0 \\
&A'_{2j-1} + v\, A'_{2j} + A'_{2j+1} = 0 \\
&\cdots\cdots\cdots\cdots \\
&A'_{2N-1} + (v + b)\, A'_{2N} = C_\beta
\end{aligned} \tag{28}$$

($2N$ Gleichungen)

Wir benutzten hier folgende Bezeichnungen:

$$\omega_h = \gamma h_0,$$

$$B_\alpha = \frac{\hbar}{2J}\frac{\omega_h \omega_{e\alpha}}{D}\left[\omega - \omega_{\beta 0} - 2(\omega_{e\alpha} + \omega_{e\beta})\right],$$

$$B_\beta = \frac{\hbar}{2J}\frac{\omega_h \omega_{e\beta}}{D}\left[\omega - \omega_{\alpha 0} - 2(\omega_{e\alpha} - \omega_{e\beta})\right],$$

[1]) Herr Dr. H. Cofta zeigte dem Autor eine während der Bearbeitung dieser Arbeit noch nicht veröffentlichte Arbeit von Pincus (1960). Seine Analyse der Resultate von Kittel (1958) ist der hier angegebenen Analyse ähnlich. Pincus geht von der Symmetrie des Hamiltonians aus und bekommt einem Ausdruck, welcher der Formel (26) analog ist.

$$C_\alpha = \frac{\hbar}{2J}\frac{\omega_h \omega_{e\alpha}}{D}\{\omega_{a\alpha}[\omega - \omega_{\beta 0} - 2(\omega_{e\alpha} + \omega_{e\beta})] + \omega_{e\beta}(\omega_{\beta 0} - \omega_{\alpha 0})\},$$

$$C_\beta = \frac{\hbar}{2J}\frac{\omega_h \omega_{e\beta}}{D}\{\omega_{a\beta}[\omega - \omega_{\alpha 0} - 2(\omega_{e\alpha} + \omega_{e\beta})] + \omega_{e\alpha}(\omega_{\alpha 0} - \omega_{\beta 0})\},$$

$$D = \omega^2 - \omega[\omega_{\alpha 0} + \omega_{\beta 0} + 2(\omega_{e\alpha} + \omega_{e\beta})] + \omega_{\alpha 0}\omega_{\beta 0} + 2\omega_{e\alpha}\omega_{\alpha 0} + 2\omega_{e\beta}\omega_{\beta 0}. \quad (29)$$

Das Glied B_α oder B_β im Ausdruck (27) hängt nicht von j ab, daher stellt es eine unendlich lange Welle dar. Das äussere Wechselfeld tritt in (28) nur im Ausdruck C_α und C_β in der ersten und letzten Gleichung auf. Dieses Feld erzeugt also Wellen endlicher Länge nur auf der Oberfläche. Ihre Amplituden A' erhält man aus (28):

$$A'_1 = \frac{-1}{D_{2N}(a,b)}\left\{C_\alpha\left[(v/u)^{\frac{1}{2}}\frac{\sin 2N\Theta}{\sin\Theta} - b\frac{\sin(2N-1)\Theta}{\sin\Theta}\right] + C_\beta\right\}$$

. . . . . .

$$A'_{2j-1} = \frac{-1}{D_{2N}(a,b)}\left\{C_\alpha\left[(v/u)^{\frac{1}{2}}\frac{\sin(2N-2j+2)\Theta}{\sin\Theta} - b\frac{\sin(2N-2j+1)\Theta}{\sin\Theta}\right] + \right.$$

$$\left. + C_\beta\left[\frac{\sin(2j-1)\Theta}{\sin\Theta} - a(v/u)^{\frac{1}{2}}\frac{\sin(2j-2)\Theta}{\sin\Theta}\right]\right\}$$

$$A'_{2j} = \frac{-1}{D_{2N}(a,b)}\left\{C_\alpha\left[\frac{\sin(2N-2j+1)\Theta}{\sin\Theta} - b(u/v)^{\frac{1}{2}}\frac{\sin(2N-2j)\Theta}{\sin\Theta}\right] + \right.$$

$$\left. + C_\beta\left[(u/v)^{\frac{1}{2}}\frac{\sin 2j\Theta}{\sin\Theta} - a\frac{\sin(2j-1)\Theta}{\sin\Theta}\right]\right\}$$

. . . . . .

$$A'_{2N} = \frac{-1}{D_{2N}(a,b)}\left\{C_\alpha + C_\beta\left[(u/v)^{\frac{1}{2}}\frac{\sin 2N\Theta}{\sin\Theta} - a\frac{\sin(2N-1)\Theta}{\sin\Theta}\right]\right\}. \quad (30)$$

Nehmen wir jetzt an, dass die Kette aus $2N+1$ Atomen besteht. Durch Anwendung der Formel (3) und (27) bekommt man für die inneren Spins die Gleichungen (28), sowie die Gleichungen

$$(u+a)A'_1 + A'_2 = C_\alpha,$$
$$A'_{2N} + (u+a)A'_{2N+1} = C_\alpha \quad (31)$$

für den ersten bzw. letzten Spinpräzessionsradius. Wenn wir jetzt die symmetrisierten A'^+ und antisymmetrisierten A'^- Amplituden, welche durch die Formeln (21) und (22) mit gestrichelten $A's$ definiert sind, anwenden, bekommen wir neue Gleichungen. Die Gleichungen für A'^- sind homogen, so dass äussere Wechselfeld in diesen nicht auftritt. Auf diese Weise stellen wir fest ,dass das äussere Wechselfeld die antisymmetri-

sche Wellen nicht erregt. Die Amplituden symmetrischer Wellen können wir mittels der Formel

$$A'_{2j-1} = \frac{C_\alpha}{a \cos N\Theta - (u/v)^{\frac{1}{2}} \cos (N+1)\,\Theta} \cos [N-(2j-2)]\,\Theta,$$

$$A'_{2j} = \frac{C_\alpha}{a\,(v/u)^{\frac{1}{2}} \cos N\Theta - \cos (N+1)\,\Theta} \cos [N-(2j-1)]\,\Theta \qquad (32)$$

ausdrücken. Der Resonanznenner in (30) und (32) ist gleich Null, wenn Θ Werte annimmt, welche die säkularen Gleichungen (6) bzw. (23) erfüllen. Wenn diese Werte komplex sind, dann gibt es eine Resonanz auch für die aperiodischen „Oberflächenspinwellen".

Wir versuchen jetzt die Tatsache aufzuklären, dass wir andere Ausdrücke für die Spinwellenamplituden erhalten, je nachdem, ob sich an den Enden der Kette gleiche oder ungleiche Spins befinden (vgl. Formel 32 und 30). Jede stehende Spinwelle im Innern des Kristalls besteht aus zwei fortlaufenden Spinwellen, die durch das äussere magnetische Wechselfeld auf seinen Begrenzungsflächen erzeugt werden. Wenn sich an beiden Enden der Kette gleiche Spins befinden, entstehen zwei gleiche Spinwellen, welche eine stehende Spinwelle mit einer Amplitude (32) ergeben. Wenn es an den Enden der Kette ungleiche Spins gibt, entstehen zwei Spinwellen mit ungleichen Amplituden und geben eine Spinwelle mit der Amplitude (30).

In dieser Arbeit haben wir nur die Austauschwechselwirkungen zwischen den nächsten Nachbaratomen berücksichtigt. Es ist bekannt, dass in den Ferriten und in den antiferromagnetischen Substanzen eine wichtige Rolle die Wechselwirkung zwischen den zweiten Nachbarn spielt. Es scheint, dass zur Lösung dieses Problems man die Rechenmethoden von Rutherford (1951) anwenden könnte, unter Berücksichtigung sowohl der Austauschwechselwirkung zwischen den erstnächsten, wie auch zwischen den zweitnächsten Nachbarn.

Die Rechenmethoden von Rutherford könnte man ebenfalls zum Problem der Spinwellenresonanz in flachen Netzen und in einfachen dreidimensionalen Gittern anwenden. Kittel (1958) zeigte, dass die in den dünnen ferromagnetischen Schichten parallel zur Oberfläche laufenden Spinwellen keinen wesentlichen Einfluss auf den Verlauf der hier beschriebenen Erscheinung ausüben. Man kann erwarten, dass dies ebenfalls bei den Ferrimagnetika und Antiferromagnetika der Fall sein wird. Das wird wahrscheinlich der Grund sein, weshalb Cofta (1960 a, b), Orbach und Pincus (1959) sich auf die lineare Kette beschränken konnten.

Herrn Prof. Dr. S. Szczeniowski danke ich für die Anregung zu dieser Arbeit, seine kritischen Hinweise und für das gefällige Durchlesen des Manuskriptes. Herrn Dr. H. Cofta danke ich für die Einführung in das Problem der halbklassischen Spinwellentheorie und seine noch nicht veröffentlichten Arbeiten, die er mir gern zur Verfügung stellte, Herrn Dr. L. Kowalewski und Herrn L. Nowicki für nützliche Diskussion.

LITERATURVERZEICHNIS

Born, M. and Kun, Huang, *Dynamical Theory of Crystal Lattices*, Oxford, Clarendon Press 1954.

Cofta, H., *Acta phys. Polon.*, **19**, 187 (1960*a*), *Acta phys. Polon.*, **19**, 193 (1960*b*).

Fox, L., *The NumericalSolution of Two-Point Boundary Problems in Ordinary Differential Equations*, Oxford, Clarendon Press 1957.

Kittel, C., *Phys. Rev.*, **110**, 1295 (1958).

Orbach, R. and Pincus, P., *Phys. Rev.*, **113**, 1213 (1959).

Pincus, P., *Phys. Rev.*, **118**, 658 (1960).

Rutherford, D. E., *Proc. Roy. Soc.* [Edinburgh], **A62**, 229 (1947). *Proc. Roy. Soc.* [Edinburgh], **A63**, 232 (1951).

Wallis, R. F., *Phys. Rev.*, **105**, 540 (1957).

THE GENERALIZATION OF DIRAC'S EQUATION. I

By Jerzy Lukierski

Institute of Theoretical Physics, University of Wrocław

(Received January 16, 1960)

An equation, invariant with respect to the 12-parameter $\boldsymbol{C}\times\boldsymbol{C}'$ group, for a spin 1/2 field interacting with an electromagnetic field is investigated. The groups $\boldsymbol{C}$ and $\boldsymbol{C}'$ are treated symmetrically. Correspondence with Dirac's equation and bilocal theory is disscused. We obtain the result that mass is an isovector, in a similar way as in Rayski's generalization of Pais' theory, and charge is an antisymmetrical isotensor of second rank. The charge and mass conjugation and the existence of a charge gauge group are investigated.

Introduction

In Dirac's equation

$$\hat{D}\psi = (\gamma_\mu \partial_\mu - m)\,\psi = 0 \tag{0.1}$$

the mass m is treated as some untransformable quantity. Thus we conclude that the groups of continuous transformations, permitted by (0.1), are

$$\boldsymbol{C}: \quad \psi' = e^{i\gamma_\mu\gamma_\nu a_{\mu\nu}}\,\psi \qquad \partial'_\mu = a_{\mu\nu}\,\partial_\nu \tag{0.2}$$

and a certain group $\boldsymbol{D}$, which commutes with $\boldsymbol{C}$ and $\hat{D}$:

$$\boldsymbol{D}: \quad \psi'' = a\psi + c\hat{C}\psi \tag{0.3}$$

where $\hat{C}$ denotes the charge conjugation operator.

We define as isotransformations with respect to a given equation these transformations, which are permitted by this equation and which commute with the proper Lorenz group $\boldsymbol{C}$. From this definition it is seen that (0.3) is the isotransformation group connected with equation (0.1).

If we introduce in (0.1) the electromagnetic field A_μ, we obtain

$$\{\gamma_\mu (\partial_\mu - ie\,A_\mu) - m\}\,\psi = 0 \tag{0.4}$$

From the assumption that charge is a scalar-isoscalar quantity it follows that the group of isotransformations reduces for (0.4) to

$$\psi'' = a\psi \tag{0.5}$$

For $|a| = 1$ (0.5) is the well-known gauge transformation group.

In this paper we enlarge the isogroups, which are permitted by equations (0.1) and (0.4)

It is well-known that the 4-spinors admit a second (besides $\boldsymbol{C}$) 6-parameter group $\boldsymbol{C}'$ (Gursey 1958, Rzewuski 1958). This group commutes with $\boldsymbol{C}$ and can be simply written as follows (Rzewuski 1958):

$$\boldsymbol{C}': \qquad \begin{aligned} c'_\alpha &= ac_\alpha + bc^*_{\dot\alpha} \\ c'^*_{\dot\alpha} &= cc + dc^*_{\dot\alpha} \end{aligned} \qquad ad - bc = 1 \tag{0.6}$$

where c_α, $c_{\dot\alpha}$ are Weyl's 2-spinors.

For Dirac's 4-spinors $\mathbf{C}'$ can be expressed in a form independent from the choice of the γ_μ-matrices representation (Tuokoka 1959):

$$\boldsymbol{C}': \quad \psi' = (A + B\gamma_5)\,\psi + (C + D\gamma_5)\,C\psi \tag{0.7}$$

$$|A|^2 - |B|^2 - |C|^2 + |D|^2 = 1 \qquad AB^* = CD^* \tag{0.8}$$

The group (0.3) restricted by the condition

$$|a|^2 - |b|^2 = 1$$

is a subgroup of $\boldsymbol{C}'$.

The product $\boldsymbol{C} \times \boldsymbol{C}'$ was investigated by Gursey (1958) and Rzewuski (1958). In Gursey's paper this group is applied to the nucleon equation. He shows that the masses of the proton and neutron remain scalar-isoscalar. We apply $\boldsymbol{C} \times \boldsymbol{C}'$ to Dirac's equation (0.1) and (0.4). It is easy to show (see Sect. 1 and Sect. 2) that mass m and charge e must be replaced by the operators $\hat{M}$ and $\hat{E}$, which transform with respect to $\boldsymbol{C}'$ as an isovector and an antisymmetrical isotensor of second rank. Equations (0.4) appear then as the general equation containing $\hat{M}$ and $\hat{E}$ written in a special isorepresentation of the mass isovector and charge isotensor.

In the general equation, in the case $A_\mu = 0$, the mass isovector components $\varkappa_{s\mu}$ are included symmetrically with the momentum operators $i\partial_\mu$ (see Sect. 6). This symmetry is most clear if we go over to the corresponding second rank equation:

$$(\Box - \varkappa^2_{s\mu})\,\psi = 0 \tag{0.9}$$

It is seen that the momentum in the usual space and the mass vector components in the isospace appear in (0.9) symmetrically. After some restriction, which can be simply interpreted, we obtain from (0.9) the generalized Pais equation, introduced by Rayski (1956).

The charge conjugation cannot be realized by means of $\boldsymbol{C}'$ (cf. (0.8)). It is obtained as a product of two isoinversions. In a special isorepresentation of our equation, corresponding to Dirac's equation, the first isoaxis and isotime are inversed.

If we consider the 4-spinors as operators, which satisfy the conventional anticommutation relations, these relations are maintained only under the 3-parametric subgroup of (0.7):

$$\psi' = a\psi + d\gamma_5 \hat{C}\psi \quad |a|^2 + |d|^2 = 1 \tag{0.10}$$

It is the well-known Pauli group, which describes isorotations in a (1,2,3)-subspace of the full Minkowski' isospace $\boldsymbol{C}'$ (Pauli 1957).

1. The mass as isovector

It is well-known that we can treat a 4-spinor as a 2×2 matrix ψ (Gursey 1956). For Weyl's representation we obtain

$$\psi_{\alpha;1} \equiv C_\alpha \qquad \psi_{\alpha;2} \equiv C^*_{\dot{\alpha}} \tag{1.1}$$

Under a proper Lorenz group $\boldsymbol{C}$ this matrix gets multiplied on the left with a 2×2 unimodular matrix. We can multiply also on the right by 2×2 matrices. If we take these matrices to be unimodular, we obtain a 6-parametric representation $\boldsymbol{C}'$ of Minkowski's isogroup. We indicate the elements of ψ with an index before the semicolon — which belongs to $\boldsymbol{C}$, and an index behind the semicolon — which belongs to $\boldsymbol{C}'$.

The Dicac equation (0.1) can be written with help of (1.1) as follows:

$$\partial_{\alpha\dot{\beta}}\, (\Psi^{\beta;\gamma})^* = \Psi_{\alpha;\delta}\, M_{\delta\dot{\gamma}} \tag{1.2}$$

where

$$M = \begin{pmatrix} m & 0 \\ 0 & -m \end{pmatrix} \tag{1.3}$$

In Dirac's equation the mass is a scalar-isoscalar. Thus we can only take into account these transformations, which do not change the matrix M. From (1.3) we see, that Dirac equation (1.2) allows only the following subgroup of $\boldsymbol{C}'$:

$$\boldsymbol{C}'_{(3)} = \begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix} \tag{1.4}$$

If we take into account the full group $\boldsymbol{C}'$, we must treat M as some isospinor of second rank with upper indices. Thus we can generalize (1.2) as follows:

$$\partial_{\alpha\dot{\beta}}\, (\Psi^{\beta;\gamma})^* = a\Psi_{\alpha;\delta}\, M^{;\delta\dot{\gamma}} \tag{1.5}$$

where a — arbitrary complex scalar-isoscalar,

M — arbitrary 2×2 hermitean matrix.

After choosing

$$\partial_{\alpha\dot{\beta};} = (\sigma_{i;}\partial_{i;} - \partial_{0;})_{\alpha\dot{\beta};} = \sigma_{\alpha\dot{\beta};}^{\mu;}\partial_{\mu;} \tag{1.6}$$

$$M_{;\alpha\dot{\beta}} = (\sigma_{;i}\varkappa_{;i} - \varkappa_{;0})_{;\alpha\dot{\beta}} = \sigma_{;\alpha\dot{\beta}}^{;\mu}\varkappa_{;\mu} \tag{1.7}$$

where $\sigma_{i;}$, $\sigma_{;i}$ are Pauli matrices, $\varkappa_{;i}$, $\sigma_{;0}$ — real, we obtain that

$$M_{;\alpha\dot{\beta}}M^{;\beta\dot{\gamma}} = (\varkappa_{;0}^2 - \varkappa_{;i}^2)\,\delta_{;\alpha}^{\dot{\gamma}} \tag{1.8}$$

and

$$\partial^{\alpha\dot{\beta};}\partial_{\beta\dot{\gamma};} = -\Box\,\delta^{\dot{\alpha}}_{\dot{\gamma};} \tag{1.9}$$

From (1.5) and its complex conjugate

$$\partial^{\alpha\dot{\beta};}\Psi_{\beta;\gamma} = a^*(\Psi^{\alpha;\delta})^* M_{;\dot{\delta}\gamma} \tag{1.10}$$

we obtain the second order equation

$$\Box\,\Psi_{\alpha;\beta} = |a|^2(\varkappa_{;i}^2 - \varkappa_{;0}^2)\,\Psi_{\alpha;\beta} \tag{1.11}$$

Comparison with Klein-Gordon equation yields:

$$|a|^2(\varkappa_{;i}^2 - \varkappa_{;0}^2) = m^2 \tag{1.12}$$

In further considerations we put

$$|a|^2 = 1 \tag{1.13}$$

and we have

$$\varkappa_{;i}^2 - \varkappa_{;0}^2 = m^2 \tag{1.14}$$

It is seen that $\Psi_{\alpha;\beta}$ from (1.5) or (1.10) satisfies also the Klein-Gordon equation with the mass m if $M^{;\dot{\delta}\gamma}$ describes a space-like mass isovector $\varkappa_{;\mu}$, the length of which is determined by (1.14) and equals m.

The usual Dirac equation (0.1) or (1.2) follows for

$$\varkappa_{;\mu}^{D} = (0, 0, -m, 0) \tag{1.15}$$

as may be seen from (1.3) and (1.7).

It is easy to show that (1.4) or (0.3) generates rotations in the (1,2,4)-pseudoeuclidean subspace of full isospace.

2. The charge as isotensor

It is well-known that in presence of the electromagnetic field we must introduce the following change in Dirac's equation (see (0.1) → (0.4)):

$$\partial_{\mu;} \to \partial_{\mu;} - ie\,A_{\mu;} \tag{2.1}$$

Because C' transformations (0.7) do not commute with ie, we see that following change

$$\partial_{\alpha\dot{\beta};} \to \partial_{\alpha\dot{\beta};} + c\, A_{\alpha\dot{\beta};} \tag{2.2}$$

where c is scalar-isoscalar and

$$A_{\alpha\dot{\beta};} = \sigma_{\alpha\dot{\beta};}^{\mu;} A_{\mu;} \tag{2.3}$$

is not equivalent to (2.1).

A second possibility is

$$\partial_{\alpha\dot{\beta};} \to \partial_{\alpha\dot{\beta};} + A_{\alpha\dot{\beta};}\, E_{;\delta}^{;\dot{\gamma}} \tag{2.4}$$

We can write (2.4) also in the form:

$$\partial_{\alpha\dot{\beta};} \to \partial_{\alpha\dot{\beta};} + A_{\alpha\dot{\beta};}\, E_{;\delta\tau}^{\ \ \tau\dot{\gamma}} \tag{2.5}$$

Equation (1.5) has now the form invariant with respect to the full $C\times C'$ group and includes the electromagnetic field $A_{\alpha\dot{\beta};}$:

$$\partial_{\alpha\dot{\beta};} (\Psi^{\beta;\gamma})^* + A_{\alpha\dot{\beta};} (\Psi^{\beta;\delta})^*\, E_{;\delta\tau}^{\ \ \tau\dot{\gamma}} = a\, \Psi_{\alpha;\delta}\, M^{;\delta\dot{\gamma}} \tag{2.6}$$

We can express $E_{;\delta\tau}^{\ \ \tau\dot{\gamma}}$ with help of an antisymmetrical tensor $e_{;\mu}^{\ \ \nu}$:

$$E_{;\delta\tau}^{\ \ \tau\dot{\gamma}} = e_{;\mu}^{\ \ \nu}\, \sigma_{;\delta\tau}^{;\mu}\, \sigma_{;\nu}^{;\tau\dot{\gamma}} \tag{2.7}$$

If we express (2.1), which acts on the 4-spinor ψ, with help of the 2×2 matrices, we obtain:

$$E^D = e_{;\mu\nu}^{D}\, \sigma_{;\mu}\, \sigma_{;\nu} \tag{2.8}$$

where

$$(e_{;}^{D}{}_{\mu\nu}) = \begin{pmatrix} 0 & -\frac{e}{2} & 0 & 0 \\ \frac{e}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{2.9}$$

Assuming

$$e_{;\mu\nu}\, e^{;\mu\nu} = e^2 \tag{2.10}$$

we obtain (2.8) as a special representation of the general (2.7). It is Dirac's representation of the isotensor $e_{;\mu\nu}$.

From (2.7) we see that

$$E_{;\delta\tau}^{\ \ \tau\dot{\gamma}} = (ie_{;i}\, \sigma_{;i} + d_{;i}\, \sigma_{;i})_{;\delta}^{\ \ \dot{\gamma}} \tag{2.11}$$

where

$$\tfrac{1}{2}\, e_{;i} = (e_{;23},\ -\, e_{;13},\ c_{;12}) \tag{2.12}$$
$$\tfrac{1}{2}\, d_{;i} = (e_{;10},\ e_{;20},\ e_{;30})$$

Condition (2.10) can be written as

$$e^2_{;i} - d^2_{;i} = e^2 \tag{2.13}$$

For the unitary subgroup (0.10) of C' $c^{;\alpha}$ has the same trans formation properties as $c_{;\alpha}$. Thus for rotations in (1,2,3)-isosubspace $e_{;i}$ is a pseudoisovector and $d_{;i}$ — an isovector. For Dirac's case we have:

$$e^D_{;i} = (0,\ 0,\ -\, e) \qquad d^D_{;i} = 0 \tag{2.14}$$

We see that in the isospace corresponding to $\boldsymbol{C}'$ we have a similar situation as in usual space. If we write the following analogy:

$$e_{;i} \sim H_{i;} \qquad d_{;i} \sim E_{i;} \tag{2.15}$$

where $E_{l;}$ and $H_{;l}$ are components of electric and magnetic strenght vectors, we obtain the following correspondence:

$$e_{;\mu\nu} \sim f_{\mu\nu;} = \partial_{\mu;} A_{\nu;} - \partial_{\nu;} A_{\mu;} \tag{2.16}$$

These analogies can be useful perhaps if we want to write in $\boldsymbol{C}'$ some equations of motion.

3. Charge conjugation

It is interesting to mention that charge conjugation is not included in (0.7). Charge conjugation $\hat{C}$ can be expressed in Weyl's 2-spinors:

$$\begin{matrix} c_\alpha \to \varepsilon c^*_{\dot\alpha} \\ c^{\cdot}_{\dot\alpha} \to \varepsilon c^*_\alpha \end{matrix} \qquad |\varepsilon| = 1 \tag{3.1}$$

From (1.1) we see that

$$\hat{C}\Psi = \Psi\, \sigma_1 \cdot e^{i\lambda\sigma_2} \tag{3.2}$$

where $\varepsilon = e^{i\lambda}$.

If we transform a charge isotensor (2.7) by means of (3.2) in virtue of (2.11) we obtain that for $\lambda = 0$

$$\begin{matrix} e_{;1} \to e_{;1} & e_{;2} \to -\, e_{;2} & e_{;3} \to -\, e_{;3} \\ d_{;1} \to d_{;1} & d_{;2} \to -\, d_{;2} & d_{;3} \to -\, d_{;3} \end{matrix} \tag{3.3}$$

Comparing (3.3) with (2.12) it is seen that (for $\lambda = 0$):

$$\hat{C} = I'_1 \cdot I'_4 \tag{3.4}$$

where I'_μ denotes inversion of the μ-th isoaxis in $\boldsymbol{C}'$.

Since for Dirac's case $e_{;1} = d_{;1} = 0$, charge conjugation changes the sign of e in (0.3).

4. *Description of equations* (1.5) *and* (2.6) *in 4-spinor notation*

Put

$$\psi = \begin{pmatrix} c_\alpha \\ c^{\dot\alpha} \end{pmatrix} = \begin{pmatrix} \Psi_{\alpha;1} \\ \Psi^{\alpha;*}_{2} \end{pmatrix} \tag{4.1}$$

From (3.1) one obtains that (for $\lambda = 0$)

$$\hat C\psi = \psi^c = \begin{pmatrix} \Psi_{\alpha;2} \\ \Psi^{\alpha;*}_{1} \end{pmatrix} \tag{4.2}$$

Using (4.1), (4.2), (1.6) and (1.7), we can write (1.5) in the following form:

$$(\gamma_\mu \partial_\mu + \hat M)\psi = 0 \tag{4.3}$$

where

$$\hat M = \varkappa_{;3} - \varkappa_{;0}\gamma_5 - (\varkappa_{;1} + i\varkappa_{;2})\gamma_5 \hat C \tag{4.4}$$

and

$$\gamma_l = \begin{pmatrix} 0 & -a^* \sigma_i \\ -a\,\sigma_i & 0 \end{pmatrix} \qquad \gamma_4 = \begin{pmatrix} 0 & -ia^* \\ ia & 0 \end{pmatrix} \tag{4.5}$$

Thus instead of the usual Dirac equation we obtain an analogous equation, where $\hat M$ is an isovector (4.4) invariant with respect to Lorentz transformations (0.2).

Equation (4.3) can be written in an another form if we introduce the following 8-component spinor Φ:

$$\Phi = \begin{pmatrix} \psi \\ \gamma_5 \psi^c \end{pmatrix} \tag{4.6}$$

In terms of Φ equation (4.3) takes the form

$$(\Gamma_\mu \partial_\mu + \vec\tau\,\vec\varkappa)\,\Phi = \varkappa_{;0}\,\Gamma_5\,\Phi \tag{4.7}$$

where

$$\Gamma_\mu = \gamma_\mu \times I^{(2)} \qquad \tau_i = I^{(4)} \times \sigma_i \tag{4.8}$$

$I^{(n)}$ — unit matrix of n-th rank,

$\vec\varkappa$ — 3-component mass-isovector in (1,2,3)-subspace of C'.

Now we introduce the general charge isotensor (2.7). Calculating with help of (4.1) and (4.2) expression (2.7) we get instead of (2.1)

$$\partial_\mu \to \partial_\mu + \hat E A_\mu \tag{4.9}$$

where

$$\tfrac{1}{2}\hat E = ie_{;12} + \gamma_5 \hat C e_{;13} + i\gamma_5 \hat C e_{;23} + \hat C e_{;14} + i\hat C e_{;24} + \gamma_5 e_{;34} \tag{4.10}$$

If we insert in (4.10) Dirac's isorepresentation (2.9) of $e_{;\mu\nu}$, we obtain the usual formula (2.1).

By virtue of (1.6), (1.7), (4.1), (4.2) and (4.5) equation, (2.6) can be written in the following form:

$$\{\gamma_\mu (\partial_\mu + \hat{E} A_\mu) + \hat{M}\} \psi = 0 \tag{4.11}$$

where $\hat{E}$ and $\hat{M}$ are determined by (4.4) and (4.10) with additional conditions (1.14) and (2.10).

Equation (4.11) may be written with help of the 8-spinor Φ. After simple calculations we obtain

$$\{\Gamma_\mu [\partial_\mu + i (\vec{e} - i \Gamma_5 \vec{d}) \vec{\tau} \cdot A_\mu] + \tau \cdot \vec{\varkappa}\} \Phi = \varkappa_{;0} \Gamma_5 \Phi \tag{4.12}$$

with

$$\vec{e}^2 - \vec{d}^2 = e^2$$
$$\varkappa_{;0} = \sqrt{\vec{\varkappa}^2 - m^2} \tag{4.13}$$

In Dirac's case we have:

$$\{\Gamma_\mu (\partial_\mu - ie\tau_3 A_\mu) - m\tau_3\} \Phi^D = 0 \tag{4.14}$$

5. *Correspondence of* (4.12) *to Dirac's equation*

Let us write (4.12) in a particularly simple isorepresentation. From (4.13) it is seen that $(\vec{e}, \vec{d})$ is a six-isovector with positive invariant form. Because Φ transform with respect to $\boldsymbol{C}'$ according to

$$\Phi' = e^{i(\vec{\alpha} + i\Gamma_5 \vec{\beta})\vec{\tau}} \Phi \qquad \vec{\alpha}, \vec{\beta}\text{-real} \tag{5.1}$$

α and β can be chosen in a such way that

$$\boldsymbol{C}' (\vec{e} - i \Gamma_5 \vec{d}) \vec{\tau} \boldsymbol{C}'^{-1} = - e \tau_3 \tag{5.2}$$

In this isorepresentation (4.12) takes the form

$$\{\Gamma_\mu (\partial_\mu - ie \tau_3 A_\mu) + \vec{\tau} \cdot \vec{\varkappa}'\} \Phi' = \varkappa'_{;0} \Gamma_5 \Phi' \tag{5.3}$$

where

$$\boldsymbol{C}' (\vec{\alpha}, -\vec{\beta}) \{\vec{\tau} \cdot \vec{\varkappa} - \varkappa_{;0} \Gamma_5\} (\boldsymbol{C}' (\vec{\alpha}, \vec{\beta}))^{-1} = \vec{\tau} \cdot \vec{\varkappa}' - \varkappa'_{;0} \Gamma_5 \tag{5.4}$$

Since $\varkappa_{;\mu}$ must be always space-like, we can take

$$\Phi'' = e^{\beta/2 \Gamma_5 \tau_3} \Phi' \tag{5.5}$$

where

$$\varkappa'_{;3} - \varkappa'_{;0} \Gamma_5 \tau_3 = \sqrt{\varkappa'^2_{;3} - \varkappa'^2_{;0}}\ e^{\beta \Gamma_5 \tau_3} \tag{5.6}$$

This pseudoeuclidean rotation transform (5.3) into

$$\{\Gamma_\mu (\partial_\mu - ie\,\tau_3\,A_\mu) + \vec{\tau}\cdot\vec{\varkappa}''\}\,\Phi'' = 0 \tag{5.7}$$

where

$$\begin{aligned} \varkappa''_{;1} &= \varkappa'_{;1} \\ \varkappa''_{;2} &= \varkappa'_{;2} \\ \varkappa''_{;3} &= \sqrt{\varkappa'^2_{;3} - \varkappa'^2_{;0}} \end{aligned} \tag{5.8}$$

After a second rotation

$$\Phi''' = e^{i\frac{\alpha}{2}\tau_3}\,\Phi'' \qquad \operatorname{tg}\alpha = \frac{\varkappa''_{;1}}{\varkappa''_{;2}} \tag{5.9}$$

which does not change the charge operator, we obtain finally:

$$\varkappa'''_{;1} = 0 \quad \varkappa'''_{;2} = \sqrt{\varkappa''^2_{;1} + \varkappa''^2_{;2}} \quad \varkappa'''_{;3} = \varkappa''_{;3} \tag{5.10}$$

Equation (4.12) can now be written in the following most simple form:

$$\{\Gamma_\mu (\partial_\mu - ie\,\tau_3\,A_\mu) + \tau_2\,\varkappa_{;2} + \tau_3\,\varkappa_{;3}\}\,\Phi = 0 \tag{5.11}$$

It is seen, that equation (5.11) is essentially different from (4.14); In Dirac's case $\varkappa_{;2} \equiv 0$. It is seen further that in the general case the mass-isovector $\vec{\varkappa}$ is not parallel to the charge-pseudoisovector $\vec{e}$.

6. *Symmetry properties of equation* (4.7)

If we take in (4.8) the Weyl representation of the γ -matrices$_\mu$

$$\gamma_i = \varrho_2\,\Sigma_i \qquad \gamma_4 = \varrho_1 \tag{6.1}$$

we obtain from (0.2) the following transformation law of the 8-spinor Φ under C-rotations:

$$\Phi' = e^{i(\vec{\Theta}\vec{A} + i\Gamma_5\vec{\Theta}\vec{B})}\,\Phi \tag{6.2}$$

Here $\vec{A}$, $\vec{B}$ — two arbitrary real 3-vectors, and

$$\vec{\Theta} = \vec{\sigma} \times I^{(4)} \tag{6.3}$$

Comparing (6.2) with (5.1) we see that transformation K:

$$\Phi \to \overline{\Phi} = \tfrac{1}{2}\,(1 + \vec{\Theta}\vec{\tau})\,\Phi = K\Phi \tag{6.4}$$

has the following property

$$K\,C\,K^{-1} = C'$$

and

$$K C' K^{-1} = C \tag{6.5}$$

because of $K^{-1} = K$.

With help of (6.1) equation (4.7) can be writen as follows:

$$\{\varrho_2 (\vec{\Sigma}\vec{\Delta} - \gamma_5 \partial_0) \times I^{(2)} + \vec{\tau}\vec{\varkappa} - \Gamma_5 \varkappa_{;0}\} \Phi = 0 \tag{6.6}$$

Performing here transformation (6.3) we obtain

$$\{\varrho_2 \times I^{(2)} \cdot (\vec{\tau}\vec{\nabla} - \Gamma_5 \partial_0) + (\vec{\Sigma}\vec{\varkappa} - \gamma_5 \varkappa_{;0}) \times I^{(2)}\} \bar{\Phi} = 0 \tag{6.7}$$

or, introducing the Γ_μ — matrices

$$(\Gamma_\mu \varkappa_{;\mu} + \vec{\tau}\vec{\nabla}) \bar{\Phi} = \Gamma_5 \partial_0 \bar{\Phi} \tag{6.8}$$

It is seen that in (4.7) the roles of the matrices Γ_μ and $(\tau_i, {}^i\Gamma_5)$ are interchanged. Formally we obtain (6.8) from (4.7) through the following substitution:

$$\varkappa_{;\mu} \rightleftarrows \partial_\mu; \tag{6.9}$$

This is connected with the fact that the choice of C as space-time rotation group and C' as isospace rotation group can be reversed.

If Φ describes plane waves with momentum p_μ; in usual space-time, the correspondence (6.9) may be written

$$\varkappa_{;\mu} \rightleftarrows i p_\mu; \tag{6.10}$$

Equation (4.7) or (6.8) describes plane isowaves in (1, 2, 3) — isosubspace. The isodirection and isovelocity of these waves are determined by the isomomentum $\varkappa_{;\mu}$. Because $\varkappa_{;\mu}$ is isospatial (see (1.14)), the isovelocity must be smaller as on the iso-lightcone. For example, Dirac's equation is an equation for standing isowaves directed paralell to the third isoaxis.

7. *Charge gauge group*

Let us investigate the general equation (4.12) written in the isorepresentation (5.11). It is seen that the following two 1-parameter groups of isorotations do not change the charge operator

$$\begin{aligned} G_1&: \ \Phi \to \Phi' = e^{i\alpha\tau_3} \Phi \\ G_2&: \ \Phi \to \Phi' = e^{\beta\Gamma_5\tau_3} \Phi \end{aligned} \tag{7.1}$$

G_1 describes rotations in the (1,2)-isoplane, G_2 — pseudoeuclidean rotations in the (3,4)-isoplane.

The group of transformations, which do not change the mass parameters, is given by the 3-parameter group M of rotations in the 3-dimensional isosubspace,

perpendicular to the mass isovector $\vec{\varkappa}\,(0, \varkappa_{;2}, \varkappa_{;3})$ from (5.11). We have

$$\boldsymbol{M}: \quad \Phi \to \Phi' = e^{T_3(\alpha_1 T_1 + \alpha_2 T_2) + i\alpha_3 T_3}\,\Phi \tag{7.2}$$

where T_i indicates the new basis:

$$T_i = \left(\tau_1, \frac{1}{m}\,(x_{;3}\tau_2 - \varkappa_{;2}\tau_3), \frac{1}{m}\,(\varkappa_{;2}\tau_2 + \varkappa_{;3}\tau_3)\right) \tag{7.3}$$

T_1 and T_2 are orthogonal to the mass vector, T_3 is to this vector parallel. Generally we can express this new basis with help of the following expressions:

$$\begin{aligned} T_1 &= \frac{1}{m^2}\,\{(\vec{\varkappa} \times \vec{a}) \times \vec{\varkappa}\} \cdot \vec{\tau} \\ T_2 &= \frac{1}{m}\,(\vec{\varkappa} \times \vec{a}) \cdot \vec{\tau} \\ T_3 &= \frac{\vec{\varkappa} \cdot \vec{\tau}}{m} \end{aligned} \tag{7.4}$$

where $\vec{a}$ is an arbitrary unit 3-vector, not parallel with $\varkappa$. To obtain the mass-isovector occuring in (5.11) we put $\vec{a} = (1, 0, 0)$.

The charge gauge transformation group is a continuous isogroup which does not change the mass and charge operators. It is given, therefore, by the common part of $\boldsymbol{G}_I$ and $\boldsymbol{M}$. It is easily seen, that only in the two cases $\varkappa_{;2} \equiv 0$ or $\varkappa_{;3} \equiv 0$ we can define continuous charge gauge transformations.

In the case $\varkappa_{;2} \equiv 0$ (Dirac's case) we have

$$T_3 = \tau_3 \tag{7.5}$$

and from (7.1) we see that $\boldsymbol{G}_1 \equiv \boldsymbol{M}$ if

$$\alpha_1 = \alpha_2 = 0 \qquad \alpha_3 = \alpha \tag{7.6}$$

$\boldsymbol{G}_1$ is the conventional gauge transformation (0.5).

In the second case, if

$$\varkappa_{;\mu} = (0, m, 0, 0) \tag{7.7}$$

we have

$$T_2 = -\tau_3$$

Taking

$$\alpha_1 = \alpha_3 = 0 \qquad \alpha_2 = -\beta \tag{7.8}$$

we obtain $\boldsymbol{G}_2 \equiv \boldsymbol{M}$.

In the general case we cannot define the 1-parameter gauge transformations. Thus, from the point of view that a continuous gauge group is obligatory, we obtain only a second possibility beyond the Dirac case. This second equation can be written

in terms of the 4-spinor ψ as follows:

$$\gamma_\mu (\partial_\mu - ieA_\mu)\, \psi = im\gamma_5 \psi^c \tag{7.9}$$

The gauge transformation group is now

$$\psi' = e^{\gamma_5 \beta}\ \psi \tag{7.10}$$

It is seen that equation (7.9) is under (7.10) invariant.

Finally we see, that mass-isovector $\vec{\varkappa}$ and charge isopseudovector $\vec{e}$ must be parallel (eq. (0.3)) or perpendicular (eq. (7.9)).

Because in (5.11) the mass operator is invariant under isoinversion I'_{14}, I'_{14} is also the charge conjugation operator for the general equation (5.11). Simultaneous mass and charge conjugation correspond to the following isorotation:

$$\Phi' = i\tau_1 \Phi \tag{7.11}$$

This is an rotation about and angle π in the (2,3)-isoplane (the mass-vector is situated on this plane). Thus mass conjugation is realised by help of the total reflexion I'_{1234} of the isospace:

$$I'_{1234}\colon\ \Phi' = i\Gamma_5 \Phi \tag{7.12}$$

By virtue of (4.12) we obtain the mass conjugation as strong reflexion in isospace or in usual space-time.

We can restrict the 6-parameter isogroup $\boldsymbol{C}'$ to Pauli's group (0.10). Because only this 3-parametric subgroup of $\boldsymbol{C}'$ preserves the conventional anticommutation relations (Pauli 1957) we must, in quantum field theory, either change these relations or restrict the group (0.7) to (0.10).

If we put in (4.7)

$$\varkappa_{;0} \equiv 0 \tag{7.13}$$

we obtain commutability of isotransformations with the full Lorentz group (including inversions). This is connected with the fact that inversions in space-time cannot be realised without iso-inversions, for example isotime inversion. If we demand isotransformations to commute with the full Lorentz group, we obtain a second reason to restrict $\boldsymbol{C}'$ to Pauli's group. A restriction of this type excludes $\boldsymbol{G}_2$ and admits only $\boldsymbol{G}_1$ gauge transformations. Dirac's equation written in the usual isorepresentation is, therefore, the only possibility in this case.

8. The connection with non-local theories

Equation (4.7) can be written in a symmetrical way, if we introduce a new coordinate-space $\eta_{;\mu}$, corresponding to the isomomentum space

$$\varkappa_{;\mu} \to i\,\frac{\partial}{\partial \eta_{;\mu}} \tag{8.1}$$

From (4.7) one obtains in this way the following Schrödinger equation

$$i\frac{\partial \Phi}{\partial t} = H\Phi \tag{8.2}$$

where

$$H = \sum_{A=1}^{A=7} \Gamma_A \frac{\partial}{\partial X_A} \qquad \begin{array}{l} A = 1, 2, 3: X_A = x_i \\ A = 4, 5, 6, 7: X_A = \eta_\mu \end{array}$$

and

$$\{\Gamma_A, \Gamma_B\} = -2\delta_{AB} \tag{8.3}$$

We have explicitly from (4.7):

$$H = \Gamma_4 \Gamma_l \frac{\partial}{\partial x_i} + i\Gamma_4 \tau_i \frac{\partial}{\partial \eta_i} - \Gamma_4 \Gamma_5 \frac{\partial}{\partial \eta_4} \tag{8.4}$$

The second order equation, corresponding to (8.2), is

$$(\Box_x \times + \Box_\eta)\Phi(x, \eta) = 0 \tag{8.5}$$

The two spaces x_μ and η_μ are independent.

If we take into account the arguments of Section 7, we can restrict the four-vectors x_μ, η_μ to their parts in (1, 2, 3)-isosubspace. From (8.5) we obtain that

$$(p_\mu^2 + \vec{\varkappa}^2)\Phi(x, \eta; p, \vec{\varkappa}) = 0 \tag{8.6}$$

This is the equation proposed in Rayski's generalization of Pais theory (see formula (21), Rayski 1956).

The author is indebted to prof. Rzewuski, prof. Rayski and prof. Łopuszański for helpfull and valuable discussions.

REFERENCES

Gursey, F., *Nuovo Cimento*, **3**, 988 (1956).
Gursey, F., *Nuovo Cimento*, **7**, 411 (1958).
Pauli, W., *Nuovo Cimento*, **6**, 204 (1957).
Rayski, J., *Acta phys. Polon.*, **15**, 407 (1956).
Rzewuski, J., *Acta phys. Polon.*, **17**, 417 (1958).
Tokuoka, Z., *Progr. theor. Phys.*, **21**, 471 (1959).

QUENCHING OF PHOTOLUMINESCENCE OF SOLUTIONS BY NON-ABSORBING FOREING MOLECULES

By Jan Głowacki

Physics Department, Normal School, Gdańsk

(Received January 22, 1960)

The theory of luminescence quenching proposed by A. Jabłoński in 1954, has been checked experimentally for liquid and rigid solutions of fluorescein, rhodamine B and "yellowish" eosin quenched by KJ and tartaric acid. This theory with a simplified model of the centre of luminescence agrees very well with the experimental results of the measurements of photoluminescence quenching. It has been also shown that, if the molecules are ions, the probability of transition of a quencher from one group to another depends, besides the viscosity of the solution and its temperature, on the charges of the luminescent and quenching molecules and on the dielectric constant of the solution.

§ 1. *Introduction*

In the theory of quenching of photoluminescence of solutions by Jabłoński (1954) the centre is considered to consist of an excited luminescent molecule enveloped by shells consisting of monomolecular layers of the solvent, which may contain quenching molecules. The quenchers may change the shells. The probability of a transition of a quencher from one shell to another depends on the viscosity of the solution and on its temperature (Jabłoński 1956). If the solution is rigid, the quenchers remain infinitely long in their shells. The probability of quenching of an excited molecule by a quencher depends on the shell in which the quencher is situated.

The mean value (i. e., the observed value) of the relative yield of luminescence for solutions containing quenchers was calculated on the basis of a simplified model of a luminescence centre (Jabłoński 1954, 1957, 1958). Assuming that the quenching probability is different from zero only for quenching by quenchers situated within a certain active sphere, the yield for a centre containing k quenchers (mutual diffusion of quenchers and luminescent molecules neglected) is (provided the volume of the sphere is suitably chosen)

$$\frac{\eta_k}{\eta_0} = \frac{1}{1+k}, \tag{1}$$

where η_0 denotes the yield for $k = 0$. By applying for k the Smoluchowski distribution (1917)

$$P_k = e^{-\nu} \frac{\nu^k}{k!}, \tag{2}$$

the following expression was obtained

$$\frac{\eta}{\eta_0} = \sum_{k=0}^{\infty} P_k \eta_k = \frac{1 - e^{-\nu}}{\nu}, \tag{3}$$

where $\nu = nv$ (n — number of quenching molecules per unit volume, v — volume of the active sphere).

The aim of the present paper is to check experimentally the above simplified version of the theory for a case of quenching by non-absorbing foreign molecules.

§ 2. *Quenching of photoluminescence of rigid solutions*

In order to check formula (3) solutions of dyes were used. Rigid solutions of fluorescent substances were obtained by using sugar syrup of great viscosity in which fluorescein and KJ (quenching substance) were dissolved. To estimate the order of magnitude of the diffusion coefficient in such a viscous solution, the simple, approximate Herzog-Polotzky formula (1914)

$$D = \frac{6.3}{\sqrt{\mu}} \cdot \frac{1}{86400}, \tag{4}$$

was used, where μ denotes the molecular weight. The very accurate experimental results obtained by Hodges and La Mer (1948) for fluorescein in water ($D = 3.7 \times \times 10^{-6}$ cm^2/sec) agree well with those calculated by means of the Herzog-Polotzky expression ($D = 4 \times 10^{-6}$ cm^2/sec; for other solvents D is inversely proportional to the viscosity). Assuming the mean life of the excited state of the luminescent molecule of fluorescein to be $\tau \approx 5 \times 10^{-9}$ sec, we obtain for the mean diffusion displacement during τ the value 0.1 Å; the radius of the solvent molecule is $r \approx 5$ Å. Therefore, the solution of fluorescein may be considered with good approximation as a rigid solution. The results of the measurements are given in Fig. 1. It may be seen that expression (3) fits very well to the experimental results.

The most interesting seem to be the results of the measurements of the fluorescence yield of rhodamine B in „sugar glass" quenched by KJ. It is worth noting that according to Vavilov and Levshin (1926) and Pringsheim (1949, p. 448) the rhodamines seem to be the only exception, being non-phosphorescent in rigid solutions at room temperature and emitting no slow fluorescence at the temperature of liquid air.

Rigid solutions of this dye were prepared by melting sugar with a small quantity of water and subsequently pouring this syrup into a metallic ring placed upon a glass plate. After several minutes, when the solution becomes very viscous, the ring were

covered with a glass plate. In this manner almost rigid solutions were obtained. These solutions were investigated at a thickness of the layer amounting to 0.3 cm. The results of the measurements are given in Fig. 1.

In all cases the photoe ctric method with photomultiplier was used for measuring the relative yield of fluorescence. The solutions of fluorescein were excited through a blue filter and the solutions of rhodamine B — through a green filter. Fluorescence

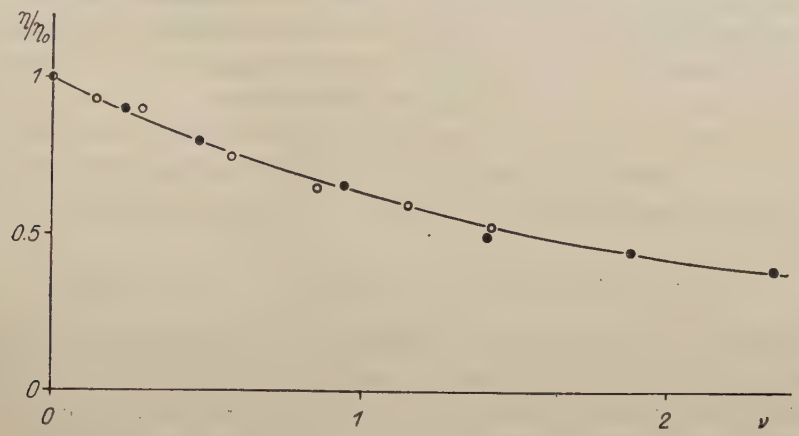

Fig. 1. Relative fluorescence yield *versus* ν.

——— — theoretical curve,
○ — expermental points for fluorescein in sugar solution quenched by KJ,
● — experimental points for rhodanine B in sugar solution quenched by KJ.

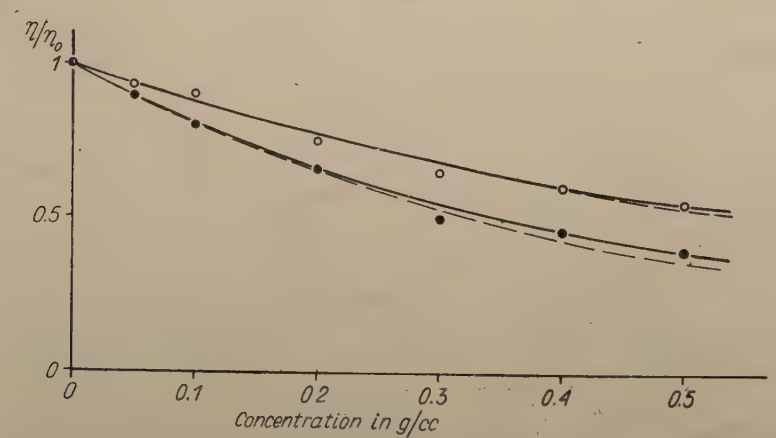

Fig. 2. Relative fluorescence yield as the function of concenctration of quenchers (KJ).

○ — experimental points for fluorescein in sugar solution quenched by KJ,
● — experimental points for rhodamine B in sugar solution quenched by KJ,
— — — theoretical curve calculated from Jabłoński's theory,
— — — — theoretical curve calculated from Frank-Vavilov's theory.

intensity for fluorescein was measured through a filter in the spectral range of 5100 to 5350 Å and for rhodamine B — through a filter in the spectral range of 5600 to 6300 Å.

It will be seen that the simplified model of the luminescent centre suffices for the description of the quenching process of fluorescence by non-absorbing molecules. In Fig. 1 the experimental values and theoretical results corresponding to $R = 5.74$ Å for fluorescein and $R = 6.8$ Å for rhodamine B are compared. The results of the measurements show that quenching is caused, most probably, by quenchers situated in the shells nearest to the luminescent molecule. Unfortunately, it was not possible to exceed the concentration of 0.5 g/cc of the quenchers.

The experimental results of quenching of rigid solutions are compared also with Frank Vavilov's theory (1931)[1]. The results of measurements show that, most probably, the quenching does not proceed according to a simple exponential law (ce, Fig. 2).

§ 3. *Quenching of photoluminescence of liquid solutions*

Although the mutual diffusion of quenchers and luminescent molecules and the orientations of their molecular axes are neglected in expression (3), this expression agrees also with the experimental results of quenching by absorbing molecules of luminescence of non-viscous solutions. In order to check expression (3) for the quenching of luminescence of liquid solutions by non-absorbing foreign molecules, aqueous solutions of fluorescein quenched by KJ were used. The results of the measurements are given in Fig. 3. Expression (3) with the radius of the active sphere $R = 17$ Å fits to the experimental results very well.

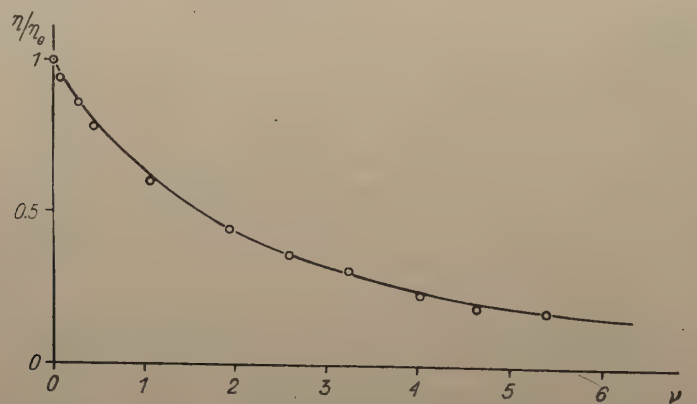

Fig. 3. Relative fluorescence yield for fluorescein in water quenched by KJ *versus* ν.

——— — theoretical curve,

○ — experimental points.

[1]) Their formula for the relative fluorescence yield is $\eta/\eta_0 = \exp(-\omega c)/(1+kc)$. If $k = 0$ (the case of rigid solution), $\eta/\eta_0 = \exp(-\omega c)$ where c is the concentration of quenchers and ω a constant.

The author investigated also the dependence of the degree of polarization of fluorescence on the concentration of the quenching molecules. This dependence is described adequately by the Perrin (1926) expression

$$\frac{1}{p} = \frac{1}{p_\infty} + \frac{\tau}{\tau_0}\left(\frac{1}{p_0} - \frac{1}{p_\infty}\right), \tag{5}$$

where p is the degree of polarization for a given concentration of quenching molecules; p_∞ — the degree of polarization for total quenching of luminescence; p_0 — the degree of polarization in the absence of quenching; τ — the mean life-time of the excited state of the luminescent molecule for a given concentration of quenchers; τ_0 — the mean life-time of the excited state of the luminescent molecule in the absence of quenching.

The well known formula given by Perrin (1929) (valid in some cases only) is

$$\frac{I}{I_0} = \frac{\eta}{\eta_0} = \frac{\tau}{\tau_0}. \tag{6}$$

From (5) and (6) we obtain the linear dependence of the reciprocal degree of polarization on the yield of luminescence

$$\frac{p_0}{p} = \frac{p_0}{p_\infty} + \frac{\eta}{\eta_0}\left(1 - \frac{p_0}{p_\infty}\right), \tag{7}$$

where

$$\frac{\eta}{\eta_0} = \frac{1 - e^{-\nu}}{\nu}.$$

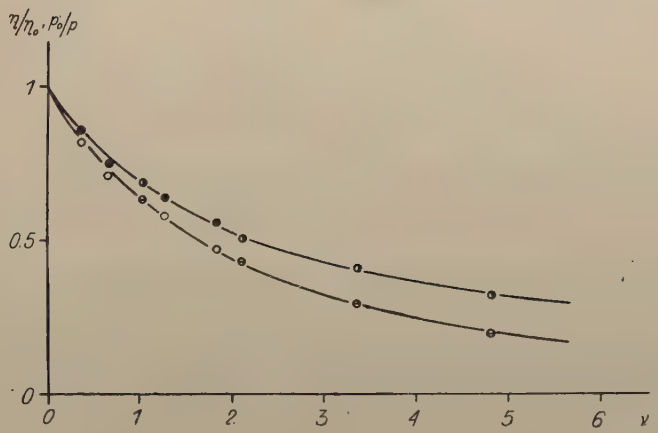

Fig. 4. Reciprocal of the degree of polarization of the fluorescence and relative fluorescence yield o rhodamine B in water quenched by KJ as the function of ν.

——— — theoretical curve,

○ and ● — experimental points of Frank-Vavilov for relative fluorescence yield and reciprocal of the degree of polarization of fluorescence,

⊖ and ◑ — author's experimental points for relative fluorescence yield and reciprocal of the degree of polarization of fluorescence.

The check of the dependence of p_0/p as function of ν is given in Fig. 4. Expression (7) was compared with the experimental results obtained by Frank and Vavilov (1931) and by the author at high concentrations of the quenchers for rhodamine B in water, quenched by KJ. The author's results of measurements of the degre of polarization of luminescence light were made by the visual method of a Savart plate and Arago compensator. The constant p_0/p_∞ was determined graphically from Graph 5.

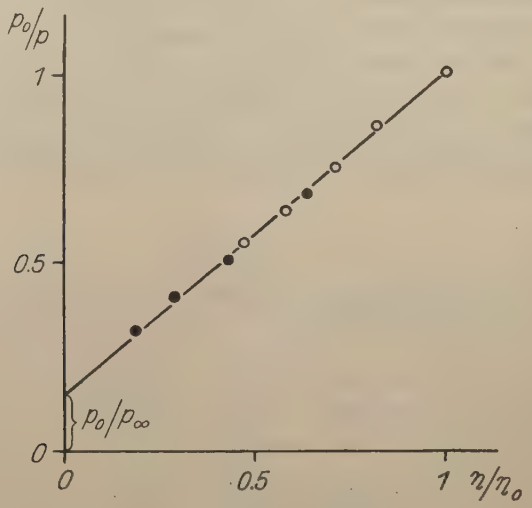

Fig. 5. p_0/p versus η/η_0.

§ 4. *Quenching of photoluminescence of water-soluble „yellowish" eosin by tartaric acid*

Another aim of the paper was to examine the quenching type of luminescence of „yellowish" eosin (microscopic dye, molecular weight 687.026) in water by tartaric acid.

Since the tartaric acid is optically active (it rotates the plane of polarization), which made measurements of the degree of polarization difficult, such investigation has not been made in this paper. The luminescence of eosin was excited with light which had passed through a Zeiss mirror monochromator giving the exciting light in the spectral region $\lambda_{max} \sim 4750$ Å. The fluorescence intensity was measured by means of a photo-multiplier (type FEU-19M) through a yellow-green filter of 5300 to 5900 Å, $\lambda_{max} = 5600$ Å.

The results of measurements of the intensity of luminescence and *pH* for different concentrations of tartaric acid are given in Figs. 6 and 7. Fig. 8 shows the absorption spectrum. The change of the absorption spectrum with the concentration of tartaric acid shows that the quenching of photoluminescence of eosin is the result of the for-

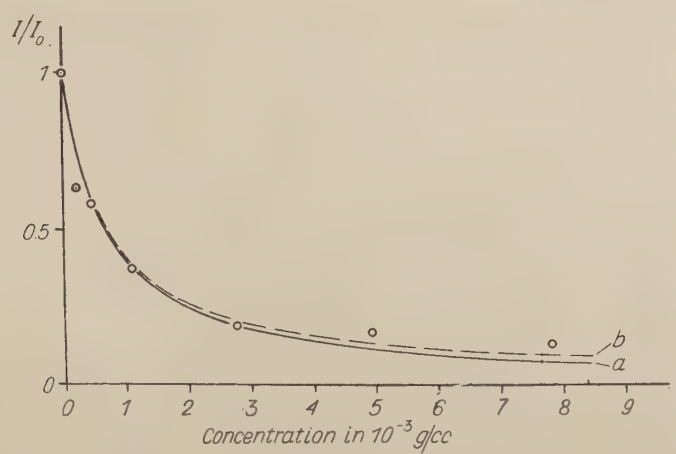

Fig. 6. Relative fluorescence intensity of eosin in water quenched by tartaric acid.
——— — tehoretical curve calculated from Eq. (8)
– – – — theoretical curve calculated from Eq. (9)
○ — experimental points.

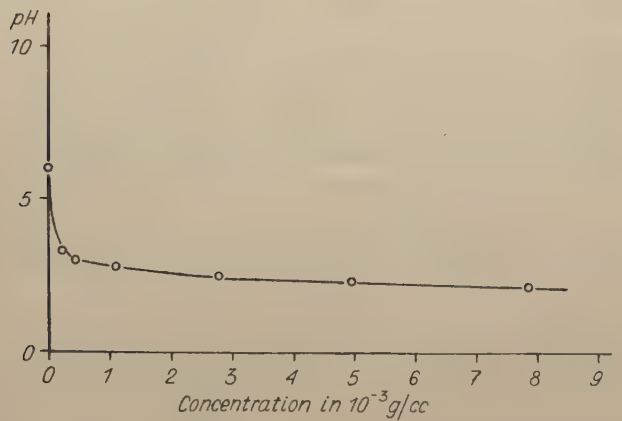

Fig. 7. *pH* of eosin solution in water versus the concentration of tartaric acid.

mation of non-fluorescent complexes of eosin and tartaric acid molecules („static" quenching). The experimental points are in good agreement with the theoretical curve of hyperbolic type (Fig. 6, curve a)

$$\frac{I}{I_0} = \frac{1}{1 + Kc}. \tag{8}$$

The insignificant systematic divergences between the experimental points and the theoretical curve at high concentration of the quenchers (tartaric acid) are probably

due to the effect of the fluorescence of complexes. Taking the fluorescence of complexes into account (Förster 1951, p. 203), we obtain a new expression

$$\frac{I}{I_0} = \frac{1 + Ac}{1 + Kc}, \tag{9}$$

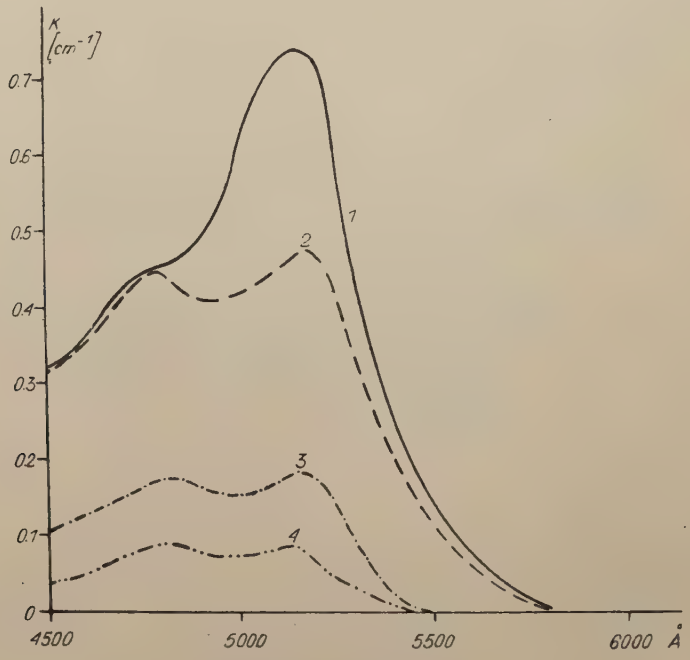

Fig. 8. Absorption spectrum of eosin in water.

1 — solution wtih c = 8.4×10^{-6} g/cc concentration of eosin without quenching molecules (tartaric acid).
2 — solution of eosin with c = 0.472×10^{-3} g/cc concentration of tartaric acid.
3 — solution of eosin with c = 1.148×10^{-3} g/cc concentration of tartaric acid.
4 — solution of eosin with c = 7.84×10^{-3} g/cc concentration of tartaric acid.

where A and K are certain constants. Expression (9) agrees better with the experimental results (Fig. 6, curve b).

Another interesting feature of the quenching of eosin by tartaric acid, studied in this paper, is a linear dependence of intensity of luminescence on the pH (Fig. 9). This dependence is given by the following empiric expression

$$\frac{I}{I_0} + C = B\,\frac{pH}{pH_0}, \tag{10}$$

where B and C are constants, I and I_0 denote the intensity of luminescence in the presence and absence of quenching, pH and pH_0 are the pH of the solution in the presence and absence of quenching respectively.

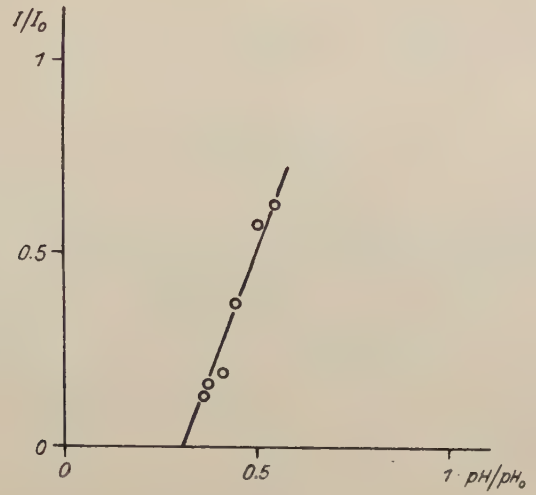

Fig. 9. I/I_0 versus pH/pH_0.

§ 5. *A note on the constant λ_{ij} for the shell model of the luminescent centre*

If the transition from one group of centres to another is taken into account, the behaviour of the system is described by a system of interdependent differential equations (Jabłoński 1954), in which case the constant λ_{ij} denotes the „transition probability" of a transition of a centre from group i to group j[2]. If we assume the simplest case in which all the centres are divided in two groups only, the centres with one single quencher in their shells forming group 1, and those with no quenchers — group 2, then λ_1 is the probability of a transition of a centre from group 1 to group 2, and λ_2 — the probability of a transition of a centre from group 2 to group 1. $\lambda_2\, dt$ denotes, therefore, the probability that a quencher enters into the active sphere of a given centre from outside during dt. It is, therefore, the process of diffusion which was investigated by Smoluchowski (1917). According to Jabłoński (1956) the following expresion was obtained

$$\lambda_2 = 4\pi DRN \tag{11}$$

by assuming that $2R\sqrt{\tau}/\sqrt{\pi D} \ll \tau$ (τ is the mean life of the excited state of the luminescent molecule), where

$$D = \frac{kT}{6\pi\eta}\left(\frac{1}{r_l} + \frac{1}{r_q}\right)$$

[2]) All the centres are divided into groups according to the occupancy of their shells by quenching molecules: the contres with one single quenching molecule in the first shell belong to group1, the centres with two quenching molecules belong to group 2, and so on.

is the diffusion coefficient of the luminescent molecule and of the quencher; R is the radius of the active sphere; N — the concentration of the quenchers. Hence (Jabłoński 1954)

$$\lambda_1 = \frac{\lambda_2}{v} = \frac{\lambda_2}{Nv}, \tag{12}$$

where v is the volume of the active sphere and $\nu = Nv$. The expression λ_2 depends, therefore, on the viscosity η of the solution, on its temperature T and on the radii of the luminescent molecule and the quencher. The expression (11) can be correct only if the fluorescent molecules carry no electric charges. If the molecules are ions, as is very frequently the case, λ_2 depends on the Coulomb electrostatic force. The diffusion problem in this case was investigated by Smoluchowski (1915), Chandrasekhar (1943) and Debye (1942). According to Debye, if the molecules are ions, $4\pi DRN$ must be multiplied by

$$A = \frac{\delta}{e^{\delta} - 1}, \tag{13}$$

where

$$\delta = \frac{Z_l Z_q e^2}{\varepsilon kTR}. \tag{14}$$

Here, $Z_l e$ and $Z_q e$ denote the charges of luminescent molecules and of quenchers, ε — the dielectric constant of the solution. If $Z_l Z_q e^2 = 0$ (one or both molecules are neutral), then $A = 1$.

By taking into account eq. (13), we obtain instead of (11)

$$\lambda_2 = 4\pi DRNA. \tag{15}$$

Expression (15) depends, therefore, on the charges of molecules and on the dielectric constant.

§ 6. *Concluding remarks*

From the examples investigated in this paper the following conclusions can be drawn. For an adequate description of luminescence quenching the simplified model of luminescence centre with the active sphere is sufficient. The experimental results of photoluminescence quenching by non-absorbing foreign molecules in liquid and rigid slsutions agree with Jabłoński's theory. It describes also adequately quenching by absorbing foreign molecules the absorption spectrum of which overlaps the fluorescence band of the luminescence molecule (Jabłoński 1957, 1958, Głowacki and Pohoski 1959).

The author is greatly indebted to Professor A. Jabłoński for suggesting the subject and his numerous friendly criticisms. He also wishes to thank Dr. J. Szychliński, Mr. T. Latowski, and Mr. Z. Polacki for their assistance with the measurements.

REFERENCES

Chandrasekhar, S., *Rev. mod. Phys.*, **15**, 1 (1943).
Debye, P., *Trans. Elektrochem. Soc.*, **82**, 265 (1942).
Förster, T., *Lumineszenz Organizcher Verbindungen*, Göttingen (1951).
Frank, I. M. and Vavilov, S. I., *Z. Phys.*, **69**, 100 (1931).
Głowacki, J. and Pohoski, R., *Bull. Acad. Polon. Sci. Cl.* III., **7**, 301 (1959).
Herzog, R. O. and Polotzky, A., *Z. phys. Chem.*, **87**, 449 (1914).
Hodges, K. C. and La Mer, V. K., *J. Amer. chem. Soc.*, **70**, 722 (1948).
Jabłoński, A., *Acta phys. Polon.*, **13**, 175 (1954).
Jabłoński, A., *Acta phys. Polon.*, **15**, 263 (1956).
Jabłoński, A., *Bull. Acad. Polon. Sci.* Cl. III., **5**, 513 (1957).
Jabłoński, A., *Bull. Acad. Polon. Sci.* Cl. III., **6**, 663 (1958).
Perrin, F., *J. Phys. Radium*, **7**, 390 (1926).
Perrin, F., *La fluorescence des solutions*, Thèse, Paris, 97 (1929).
Pringsheim, P., *Fluorescence and Phosphorescence*, Interscience Publishers Inc. New York (1949).
Smoluchowski, M., *Ann. Phys.*, **48**, 1103 (1915).
Smoluchowski, M., *Z. phys. Chem.*, **92**, 129 (1917).
Vavilov, S. I. and Levshin, V. S., *Z. Phys.*, **35**, 920 (1926).

THE UNIVERSAL FERMI INTERACTION IN THE MUON CAPTURE

By W. Majewski

Institute of Physics, Polish Academy of Sciences, Warszawa

(Received January 31, 1960)

The angular asymmetry of neutrons from the capture of polarized μ-mesons in complex nuclei is estimated, assuming the applicability of the Fermi gas model of the nucleus. It is shown that the four-fermion V—A interaction together with the conserved current hypothesis lead to considerable asymmetry.

1. The idea of universality of weak interactions leading to processes with four fermions expressed ten years ago[1] has got in the last two years strong experimental and theoretical support. The notion of the universality was made more precise. Namely, it does not mean that these processes may be described by a set of couplings with the same constants C_S, C_V, etc.; the universality is expected for a primary weak interaction, which may lead to different effective four-fermion couplings in different elementary processes as a result of an internal structure of strongly interacting particles entering in these processes. It is known that the experiments on β-and μ-decays are consistent with V-A coupling. To explain non-renormalization of the vector coupling constant in β-decay, Feynman and Gell-Mann[2] supposed that into the vector coupling with (ev)-curent enter not simply the nucleon current, but a total conserved isovector current. This hypothesis leads to additional physical consequences: it gives corrections to the energy spectrum of β-electrons[3], introduces the angular correlation β—γ in allowed transitions[4], etc. These effects are, however, relatively small because of the small nucleon momentum transfer involved in β-decay ($q^2 \equiv \equiv (p_n - p_p)^2 \approx 0$)*).

If the idea of the universality is correct, then the least studied up to now μ^--capture proces must be described by the V-A interaction with the conserved vector current. The effect of the latter — if it exists — must manifest itself here more significantly because of the larger momentum transfer ($q^2 \approx 0.9\, m^2$, m—the muon mass). In the present paper the process of muon capture is considered from the point of view of the examination of V-A coupling and especially of the conserved vector current hypothesis.

*) We adopt $\hbar = c = 1$

2. We assume

$$H_{\text{int}}^{\text{weak}} = \frac{G}{\sqrt{2}} j_\lambda^V [\bar{\psi}_\nu \gamma_\lambda (1+\gamma_5)\psi_\mu] - \frac{G}{\sqrt{2}} j_\lambda^A [\bar{\psi}_{\nu i} \gamma_\lambda \gamma_5 (1+\gamma_5)\psi_\mu] + \text{h.c.},$$

$$j_\lambda^V = \bar{\psi}_N \tau^- \gamma_\lambda \psi_N + i [\varphi_\pi^* T \nabla_\lambda \varphi_\pi - (\nabla_\lambda \varphi_\pi)^* T^- \varphi_\pi] + \text{(other posisble terms)},$$

$$j_\lambda^A = \bar{\psi}_N \tau^- i \gamma_\lambda \gamma_5 \psi_N.$$

The vector current j_λ^V fulfils the condition $j_{\lambda,\lambda}^V = 0$. To the first order in the weak interaction the S-matrix element for muon capture by a proton has the following form

$$S_{fi} = -i (2\pi)^4 \delta^4 (p_n + p_\nu - p_p - p_\mu) \mathfrak{M},$$

$$\mathfrak{M} = \frac{G}{\sqrt{2}} < n | j_\lambda^V | p > [\bar{u}_\nu \gamma_\lambda (1+\gamma_5) u_\mu] - \frac{G}{\sqrt{2}} < n | j_\lambda^A | p > [\bar{u}_\nu i \gamma_\lambda \gamma_5 (1+\gamma_5) u_\mu].$$

Here $|n>, |p>$ denote the states of physical nucleons. The effect of strong interaction is taken into account phenomenologically by writing the nucleon matrix elements in a general form following from the Lorentz-covariance. Taking into account the charge symmetry and the T-invariance of the strong interactions, we have only four form-factors[5]:

$$< n | j_\lambda^V | p > = f_V \bar{u}_n \left[\gamma_\lambda + \frac{\varkappa}{2M} \sigma_{\lambda\varrho} (p_n - p_p)_\varrho \right] u_p$$

$$< n | j_\lambda^A | p > = f_V \bar{u}_n \left[\lambda i \gamma_\lambda \gamma_5 + \frac{\xi}{m} (p_n - p_p)_\lambda \gamma_5 \right] u_p;$$

M is the nucleon mass; $u_{n,p,\nu,\mu}$, are bispinors of the corresponding particles, which are for the present assumed to be free; $f_V, \varkappa, \lambda, \xi$ are functions of q^2. Using the conservation law of the 4-momentum and the Dirac equations for leptons, we obtain

$$\mathfrak{M} = \frac{G}{\sqrt{2}} f_V \left\{ \left[\bar{u}_n \left(\gamma_\lambda + \frac{\varkappa}{2M} \sigma_{\lambda\varrho} (p_n - p_p)_\varrho \right) u_p \right] [\bar{u}_\nu \gamma_\lambda (1+\gamma_5) u_\mu] + \right.$$

$$\left. + \lambda [\bar{u}_n \gamma_\lambda \gamma_5 u_p] [\bar{u}_\nu \gamma_\lambda (1+\gamma_5) u_\mu] + \xi [\bar{u}_n \gamma_5 u_p] [\bar{u}_\nu (1-\gamma_5) u_\mu] \right\}.$$

We see that taking into account of the effect of strong interactions changes the effective V and A interaction strengths: $G/\sqrt{2} \to f_V G/\sqrt{2}$, $-G/\sqrt{2} \to -\lambda f_V G/\sqrt{2}$, and introduces two new couplings: a pseudoscalar one with a factor $c_P^{\text{eff}} = -\xi f_V G/\sqrt{2}$, and a derivative coupling — the so called weak magnetism term — proportional to

$$\frac{\varkappa}{2M} f_V \frac{G}{\sqrt{2}}.$$

The conservation of the vector current j_λ^V and the charge invariance of the strong interaction lead to the equality

$$<n|j_\lambda^V|p> = <N|j_\lambda^z|N> =$$

$$= \bar{u}(p_2)\left[\gamma_\lambda (F_1^p - F_1^n) + \frac{1}{2M}(\varkappa^p F_2^p - \varkappa^n F_2^n)\,\sigma_{\lambda\varrho}(p_2 - p_1)_\varrho\right] u(p_1).$$

j_λ^z is the isovector part of the charge-current density of nucleons and pions; $\varkappa^p = 1.79$, $\varkappa^n = -1.91$ are the anomalous magnetic moments of proton and neutron, respectively, in units $e/2M$;

$F_1^{p,n}$, $F_2^{p,n}$ are the nucleon electromagentic formfactors, known from experiment [6] to be

$$F_1^n \approx 0,\ F_1^p = F_2^p = F_2^n \approx 1 - \tfrac{1}{6} < r >^2 q^2 = 0.97$$

for the momentum transfer involved in muon capture, if the root mean square radius of the nucleon charge density is $<r> = 0.80 \times 10^{-13}$ cm. The formfactors in β-decay and μ-capture were investigated in[5] by the method of dispersion relations; the result is

$$f_V \lambda|_{q^2=0.9m^2} \approx f_V\,\lambda|_{q^2=0} \text{ and } (f_V\,\xi)/(f_V\,\lambda)|_{q^2=0.9m^2} \approx 8.$$

The formfactors f_V, etc. for $q^2 = 0$ are measured in β-decay, therefore, using the experimental values $f_V(0) \approx 1$, $\lambda(0) \approx 1.25$**), one obtains for muon capture. $f_V = 0.97$, $\varkappa = \varkappa^p - \varkappa^n = 3.7$, $\lambda = 1.29$, $\xi = 10.3$. Let us notice that if we do not assume the conservation of the vector current, then $\varkappa \approx 0.25$[5], and the weak magnetism term is negligibly small in muon capture too.

Rejecting the terms $\sim (v_{\text{nucl}}/c)^2$ and assuming the muon to be at rest at the moment of capture, we obtain

$$\mathfrak{M} = \frac{G}{\sqrt{2}} f_V \frac{1}{\sqrt{2}} (1 - <\vec{\sigma}>_L \vec{\nu})\,(a + \vec{b} <\vec{\sigma}>_N)$$

$$a = 1 + \frac{\vec{p}_p + \vec{p}_n}{2M} \cdot <\vec{\sigma}>_L$$

$$\vec{b} = -\left[\lambda\left(<\vec{\sigma}>_L + \frac{\vec{p}_p + \vec{p}_n}{2M}\right) + \xi \frac{\vec{p}_p - \vec{p}_n}{2M} + i\mu \frac{(\vec{p}_p - \vec{p}_n)}{2M} \times <\sigma>\right], \quad (1)$$

$\vec{\nu} = \vec{p}_\nu/p_\nu$; $<\vec{\sigma}>_N$, $<\vec{\sigma}>_L$, are matrix elements taken between the spins states of nucleons and leptons, respectively; $\mu = 1 + \varkappa^p - \varkappa^n$.

3. In [7] it was shown that the effect of the conserved vector current (which reduces essentially to an about fifteenfold increase of the coefficient $\varkappa$) considerably changes the angular asymmetry of the neutrons from capture of polarized muons in hydrogen. With inclusion of an effective pseudoscalar term (omitted in [7]) we obtain from (1) for capture in hydrogen ($\vec{p}_p = 0$):

$$dw = \frac{1}{(2\pi)^2} \frac{G^2}{2} f_V^2 \frac{1}{\pi a_0^3} I\,(1 + \alpha \vec{P}_\mu \vec{n})\, p_n^2\, d\Omega_n$$

$$I = 1 + 3\lambda^2 + \beta(1 + \lambda^2 + 2\lambda\mu - \lambda\xi) + \tfrac{1}{4}\beta^2(1 + \lambda^2 + 2\mu^2 + \xi^2 - 2\lambda\xi) \quad (2)$$
$$I\alpha = 1 - \lambda^2 + \beta(1 + \lambda^2 - 2\lambda\mu - \lambda\xi) + \tfrac{1}{4}\beta^2(1 + \lambda^2 - 2\mu^2 + \xi^2 - 2\lambda\xi),$$

$\vec{n} = \frac{p_n}{p_n}$ (here and further on $p_n \equiv |\vec{p}_n|$), $\beta = \frac{p_n}{M}$, $\vec{P}_\mu$ is the polarization vector of captured muons). The space part of the muon wavefunction is taken as a solution for the K-orbit at zero: $|\varphi(0)|^2 = 1/\pi a_0^3$, $a_0 = (mc^2)^{-1}$. Numerical estimation leads to the result that the effect of a conserved vector current increases the capture rate by 17%; still greater is its influence on the asymmetry parameter α: it increases by 33%, from -0.33 to -0.44. (This α is calculated neglecting the hyperfine structure of the μ-mesic atom).

As pointed out in [8], the experiment on the neutron asymmetry is practically unfeasible in hydrogen — there occurs a total depolarization of muons with transition of μ-mesic atoms into the singlet state from which the capture goes on. It follows from (1) that

$$w_{\text{singl}} \sim I - 3J,$$
$$J = -2\lambda(\lambda + 1) + \tfrac{1}{3}\beta[-4\lambda(\mu + 1) - 2\lambda^2 + \xi + 2\lambda\xi] + \tfrac{1}{6}\beta^2[\mu^2 + 2\mu + \xi].$$

It follows that the conserved vector current hypothesis increases the singlet capture rate by 11% in comparison with the usual V-A theory.

4. At present only experiments on μ^--meson capture in nuclei heavier than Be are being performed. For this reason it seems worthwhile to estimate the expec ted neutron asymmetry for capture incomplex nuclei: we are interested in the investigation of a decrease of this asymmetry in comparison with the hydrogen case. It may be emphasized that the principal contribution to α becomes from the relativistic terms, namely the pseudoscalar and weak magnetism ones (in the nonrelativistic limit one obtains from (2): $\alpha = -0.11$). We adopt for the nuclear proton the effective interaction hamiltonian yielding for the free proton the matrix element (1). Our estimation of the asymmetry will be probably only qualitative: the proton motion inside the nucleus is regarded according to Fermi gas model which already was applied previously in muon capture calculations (see e.g. [9, 10], where the validity of this model is also discussed in detail).

Namely we suppose that nucleons move independently as free particles in the sphere of radius R, being in potential well of the depth $V_0 = E_F + E_B$, and occupy all the momentum states up to P_F determined from the condition

$$2 \cdot \frac{4}{3}\pi R^3 \cdot \frac{4}{3}\pi P_F^3 \cdot \frac{1}{(2\pi)^3} = Z = N = \frac{A}{2}$$

) See e. g. J. A. Smorodynski *Uspekhi Fiz. Nauk*, **68, 653 (1959).

(for the sake of simplicity one takes $Z = N = A/2$, Z and N are numbers of protons and neutrons in the nucleus, respectively, with $A = Z + N$); $R = R_0 A^{1/3}$, $R_0 = 1.2 \times 10^{-13}$ cm, $E_F = P_F^2/2M \approx 33\ MeV$, $E_B \approx 8$ MeV — the binding energy of last neutron. The calculation consists in an integration over the momentum distribution of protons of the probability

$$dw = 2\pi |\mathfrak{M}|^2 \frac{d^3 p_n}{(2\pi)^3} \left(\frac{4}{3}\pi P_F^3\right)^{-1} \frac{A}{2} d^3 p_p / dE.$$

It may be shown that the energy — momentum conservation and the Pauli principle for neutrons impose on the proton momentum — for given neutron jkinetic energy inside the nucleus E_n — the lower limit

$$p_p \geqslant p_{\min} = [(M + \sqrt{2ME_n})^2 - 2Mm]^{1/2} - M.$$

Because of the approximate character of the applied model only principal relativistic terms were retained in the calculation (that is, terms with the coefficients μ and ξ). For the nucleus of spin 0 the following angular distribution of neutrons with respect to the muon polarization was obtained

$$dw \sim I(E_n)\,[1 + \alpha(E_n)\,\vec{P}_\mu \vec{n}]\,dE_n\,d\Omega_n$$

$$I(E)_n = \left[(1 + 3\lambda^2)\frac{k^2}{2} + \lambda(2\mu - \xi)\frac{1}{M}\frac{k^3}{3}\right]\Bigg|_{k_1}^{k_2}$$

$$I(E_n)\,\alpha(E_n) = \frac{M}{\sqrt{2ME_n}}\left[(1 - \lambda^2)\left(\frac{1}{2M}\frac{k^3}{3} - \frac{k^2}{2} + mk\right) - \right.$$

$$\left. - \lambda(2\mu + \xi)\frac{1}{M}\left(\frac{1}{2M}\frac{k^4}{4} - \frac{k^3}{3} + m\frac{k^2}{2}\right)\right]\Bigg|_{k_1}^{k_2}$$

$$k_1 = M + \sqrt{2ME_n} - [(M + \sqrt{2ME_n})^2 - 2Mm]^{1/2}$$

$$k_2 = E_F + m - E_n.$$

It follows from kinematics that the emitted neutrons have kinetic energies outside the nucleus up to about 16 MeV. $I(E_n)$ is the neutron spectrum, $\alpha(E_n)$ is the angular asymmetry coefficient depending now on E_n; both these functions are represented in Fig. 1. We see that α varies in the limits from — 0.37 for slow neutrons to — 0.45 for fast neutrons, when for the case of a proton at rest in used approximation $\alpha = -0.55$. It means that the neutron asymmetry is considerably retained for muon capture in nuclei (67—80% of that in hydrogen). Then not only an experimental observation of this asymmetry may be expected, but the conserved vector current hypothesis may be probably checked too.

5. Up to now there were reported three experimental results on neutron asymmetry in muon capture: 1) $|\alpha P_\mu| = 0.03 \pm 0.15$ [11], 2) $\alpha P_\mu = -0.05 \pm 0.02$ [12] (both for capture in sulphur), and 3) $\alpha = +0.15 \pm 0\,18$ [13] (for capture in magnesium; this

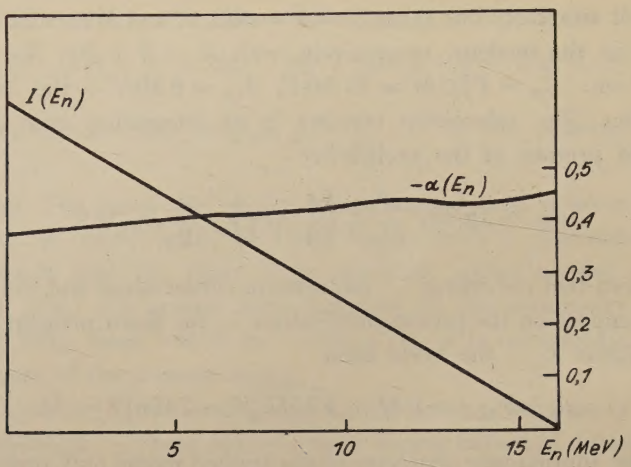

Fig. 1. I (E_n) — the neutron spectrum in arbitrary units, α (E_n) — the angular asymmetry parameter.

value was measured for emitted neutrons with $E_n > 5.5$ MeV). These experiments are difficult because of the great gamma ray and evaporation neutron background, and their accuracy is for the present unsufficient to tell anything conclusive about the μ^--capture theory adopted here.

The author wishes to express his cordial thanks to Dr. Chou Huang-Chao for collaboration and advice during his stay at the J.I.N.R.***)

REFERENCES

[1] See e. g. J. Tiomno and J. A. Wheeler, *Rev. mod. Phys.*, **21**, 153 (1949).
[2] R. Feynman and M. Gell-Mann, *Phys. Rev.*, **109**, 193 (1958).
[3] M. Gell-Mann, *Phys. Rev.*, **111**, 362 (1958).
[4] J. Bernstein and R. R. Lewis, *Phys. Rev.*, **112**, 232 (1958).
[5] M. L. Goldberger and S. B. Treiman, *Phys. Rev.*, **111** 354 (1958).
[6] R. Hofstadter, F. Bumiller and M. R. Yearian, *Rev. mod. Phys.*,**30**, 482 (1958).
[7] Chou Huang- chao and W. Majewski, *Zh. eksper. teor. Fiz.*,**35**, 1581 (1958).
[8] S. S. Gerstein, *Zh. eksper. teor. Fiz.*, **34**, 463, 993 (1958).
[9] H. Überall, *Nuovo Cimento*, **6**, 533 (1957).
[10] M. Cini and R. Gatto, *Nuovo Cimento*, **11**, 253 (1959).
[11] Coffin, Sachs, Tycko, *Bull. Am. Phys. Soc.*, **3**, 52 (1958).
[12] Astbary and others (mentioned in Alikhanov's report on the Kiev conference, 1959).
[13] W. F. Baker and C. Rubbia, *Phys. Rev. Letters*, **3**, 179 (1959).

***) The essential part of this paper was finished at the beginning of 1959 in the Joint Institute for Nuclear Research (J. I. N. R.) in Dubna (U.S.S.R.). After that time other papers have been published dealing with the same problem, see e. g. H. Primakoff, *Rev. mod. Phys.*, **31**, 802 (1959); A. Fujii and H. Primakoff, *Nuovo Cimento*, **12**, 327 (1959).

Volumen XIX — Fasciculus 4